EXHIBIT 2

3

Molecular Cloning

A LABORATORY MANUAL

SECOND EDITION

J. Sambrook
UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER

E.F. Fritsch
GENETICS INSTITUTE

T. Maniatis
HARVARD UNIVERSITY

Cold Spring Harbor Laboratory Press
1989

Molecular
Cloning
A LABORATORY MANUAL
SECOND EDITION

Printed in th United States of America

9

Book and cover design by Emily Harste

Cover: The electron micrograph of bacteriophage λ particles stained with uranyl acetate was digitized and assigned false color by computer. (Thomas R. Broker, Louise T. Chow, and James I. Garrels)

Cataloging in Publications data

Sambrook, Joseph
Molecular cloning : a laboratory manual / E.F. Fritsch, T. Maniatis—2nd ed.
p. cm.
Bibliography: p.
Includes index.
ISBN 0-87969-309-6
1. Molecular cloning—Laboratory manuals. 2. Eukaryotic cells--Laboratory manuals. I. Fritsch, Edward F. II. Maniatis, Thomas III. Title.
QH442.2.M26 1987
574.87'3224—dc19 87-35464

All Cold Spring Harbor Laboratory Press publications may be ordered directly from Cold Spring Harbor Laboratory Press, 10 Skyline Drive, Plainview, New York 11803. Phone: 1-800-843-4388. In New York (516) 367-8423. FAX: (516) 367-8432.

16

Expression of Cloned Genes in Cultured Mammalian Cells

The development of methods for the introduction of DNA into cultured mammalian cells has made it possible to express cloned genes in a broad range of cell types from different species. These methods have been used to overproduce proteins for structural and biochemical studies and to identify elements involved in the control of gene expression. In both types of studies, the cloned sequence of interest is inserted into the appropriate expression vector, cloned in bacteria, amplified by replication, and then used to transfect mammalian cells.

In this chapter, we describe a number of commonly used mammalian expression vectors, and we provide protocols for introducing cloned genes into mammalian cells. We begin by discussing expression of proteins and then go on to describe methods used to study gene regulation. In general, different vectors are required for the two types of studies, but many of the basic components used in the construction of these vectors are the same.

EXPRESSION OF PROTEINS FROM CLONED GENES

A few eukaryotic proteins have been expressed efficiently and inexpensively in prokaryotic hosts (see Chapter 17). However, many eukaryotic proteins synthesized in bacteria fold incorrectly or inefficiently and, consequently, exhibit low specific activities. In addition, production of authentic, biologically active eukaryotic proteins from cloned DNA frequently requires posttranslational modifications such as accurate disulfide bond formation, glycosylation, phosphorylation, oligomerization, or specific proteolytic cleavage—processes that are not performed by bacterial cells. This problem is particularly severe when expression of functional membrane or secretory proteins such as cell surface receptors and extracellular hormones or enzymes is required.

Because of these problems, considerable effort has been made to develop systems to express mammalian proteins in mammalian cells. These systems can be divided into two types: those that involve transient or stable expression of transfected DNA and those that involve the use of viral expression vectors derived from simian virus 40 (SV40) (Elder et al. 1981; Gething and Sambrook 1981; Rigby 1982, 1983; Doyle et al. 1985; Sambrook et al. 1986), vaccinia virus (Mackett et al. 1985; Moss 1985; Fuerst et al. 1986, 1987), adenovirus (Solnick 1981; Thummel et al. 1981, 1982, 1983; Mansour et al. 1985; Karlsson et al. 1986; Berkner 1988), retroviruses (Dick et al. 1986; Gilboa et al. 1986; Eglitis and Anderson 1988), and baculoviruses (Luckow and Summers 1988). The diversity of these animal viruses is so great that an account of their molecular biology is beyond the scope of this chapter. In addition, effective utilization of viral vectors requires some prior experience in the methods used to grow, quantitate, and plaque-purify different viruses. An entire chapter would be required to provide all of the information necessary to use each of these vectors. For these reasons, we have chosen to focus entirely on expression methods that involve DNA transfection. (*Note*: An excellent manual describing methods for the use of baculovirus vectors and procedures for culture of insect cells has been published by Summers and Smith [1987].)

Expression of proteins from cloned eukaryotic genes in mammalian cells has been used for a number of different purposes:

- To confirm the identity of a cloned gene by using immunological or functional assays to detect the encoded protein
- To express genes encoding proteins that require posttranslational modifications such as glycosylation or proteolytic processing
- To produce large amounts of proteins of biological interest that are normally available in only limited quantities from natural sources
- To study the biosynthesis and intracellular transport of proteins following their expression in various cell types

- To elucidate structure-function relationships by analyzing the properties of normal and mutant proteins
- To express intron-containing genomic sequences that cannot be transcribed correctly into mRNA in prokaryotes or yeasts
- To identify DNA sequence elements involved in control of gene expression

When choosing a mammalian expression vector, the following parameters should be taken into account:

- *The species and types of host cells that are available.* Not all types of mammalian cells can be transfected efficiently and not all of them will necessarily carry out exactly the same set of posttranslational modifications. For example, Chinese hamster ovary (CHO) cells typically add more terminal sialic acid residues to secretory and transmembrane proteins than do simian CV-1 cells or mouse NIH-3T3 cells. Whether or not this difference is significant may vary according to the nature of the protein that is expressed. In addition, some lines of cultured cells may endogenously synthesize high levels of the protein that is expressed. This can complicate functional and immunological assays.
- *Whether the experiment can be carried out with cells that transiently express the foreign protein or whether it will be necessary to isolate cell lines that permanently express the protein.* For example, immunofluorescent localization of a protein can be carried out just as well with transiently transfected cells as with stable cell lines. However, production of more than a few micrograms of foreign protein can usually be achieved only by the development of suitable stable cell lines.
- *The size of the gene that is to be transfected and expressed.* Some mammalian viral vectors have strict packaging requirements and will not accept large pieces of foreign DNA.
- *The presence of controlling elements in the transfected DNA.* Cloned cDNAs can only be expressed if they are correctly placed in a vector that supplies a promoter and other elements such as enhancers, splice acceptor and/or donor sequences, and polyadenylation signals. Genomic DNA sequences may already carry these controlling elements, but there is no guarantee that they will work normally in the lines of cultured cells that are available. This is a problem particularly when dealing with genes that are expressed in a tissue-specific fashion.

Plasmid vectors that have been used to introduce and express cloned genes in mammalian cells can be divided into three major classes:

1. Simple plasmid-based vectors that contain no eukaryotic replicon.
2. More complex plasmid vectors that incorporate elements from the genomes of eukaryotic viruses to increase the copy number of the transfected DNA and the efficiency with which foreign proteins are expressed.
3. Vectors designed to facilitate amplification of transfected sequences that become integrated into the host genome.

FUNCTIONAL COMPONENTS OF MAMMALIAN EXPRESSION VECTORS

Mammalian expression vectors contain both prokaryotic sequences that facilitate the propagation of the vector in bacteria and one or more eukaryotic transcription units that are expressed only in eukaryotic cells. The eukaryotic transcription unit consists of noncoding sequences and sequences coding for selectable markers. It is frequently assembled as a composite of elements derived from different, well-characterized viral or mammalian genes. The components that are used in various expression vectors are described briefly below.

Prokaryotic Plasmid Sequences That Facilitate the Construction, Propagation, and Amplification of Recombinant Vector Sequences in Bacteria

The essential prokaryotic elements include a replicon that functions in *Escherichia coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor the recombinant plasmids, and a limited number of unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. Most of the mammalian vectors in current use contain prokaryotic sequences from derivatives of the plasmid pBR322 (e.g., pXf3, pBRd, and pML) that lack sequences that seem to interfere with the replication of the transfected DNA in eukaryotic cells (Lusky and Botchan 1981). Deletion of unnecessary segments of plasmid DNA also reduces the size of the vector and facilitates the placing of unique restriction sites that can be utilized for the insertion and manipulation of eukaryotic sequences.

A Eukaryotic Expression Module That Contains All of the Elements Required for the Expression of Foreign DNA Sequences in Eukaryotic Cells

The most basic eukaryotic expression module contains a promoter element to mediate transcription of foreign DNA sequences and signals required for efficient polyadenylation of the transcript. Additional elements of the module may include enhancers and introns with functional splice donor and acceptor sites.

PROMOTER AND ENHANCER ELEMENTS

Unlike the signals required for RNA processing, which function efficiently in all types of mammalian cells, the activities of elements that control transcription—promoters and enhancers—vary considerably among different cell types. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription (for review, see Dynan and Tjian 1985; Serfling et al. 1985; McKnight and Tjian 1986; Sassone-Corsi and Borrelli 1986; Maniatis et al. 1987). The combination of different recognition sequences and the amounts of the cognate transcription factors determine the efficiency with which a given gene is transcribed in a particular cell type.

Many eukaryotic promoters contain two types of recognition sequences: the

TATA box and the *upstream promoter elements*. The TATA box, located 25–30 bp upstream of the transcription initiation site, is thought to be involved in directing RNA polymerase II to begin RNA synthesis at the correct site. In contrast, the upstream promoter elements determine the rate at which transcription is initiated. These elements can act regardless of their orientation, but they must be located within 100 to 200 bp upstream of the TATA box. *Enhancer elements* can stimulate transcription up to 1000-fold from linked homologous or heterologous promoters. However, unlike upstream promoter elements, enhancers are active when placed downstream from the transcription initiation site or at considerable distances from the promoter. Many enhancers of cellular genes work exclusively in a particular tissue or cell type (for review, see Voss et al. 1986; Maniatis et al. 1987). In addition, some enhancers become active only under specific conditions that are generated by the presence of an inducer, such as a hormone or metal ion (for review, see Sassone-Corsi and Borrelli 1986; Maniatis et al. 1987). Because of these differences in the specificities of cellular enhancers, the choice of promoter and enhancer elements to be incorporated into a eukaryotic expression vector will be determined by the cell type(s) in which the recombinant gene is to be expressed. Conversely, the use of a prefabricated vector containing a specific promoter and cellular enhancer may severely limit the cell types in which expression can be obtained.

Many enhancer elements derived from viruses have a broader host range and are active in a variety of tissues, although significant quantitative differences are observed among different cell types. For example, the SV40 early gene enhancer is promiscuously active in many cell types derived from a variety of mammalian species, and vectors incorporating this enhancer have consequently been widely used (Dijkema et al. 1985). Two other enhancer/promoter combinations that are active in a broad range of cells are derived from the long terminal repeat (LTR) of the Rous sarcoma virus genome (Gorman et al. 1982b) and from human cytomegalovirus (Boshart et al. 1985).

TERMINATION AND POLYADENYLATION SIGNALS

During the expression of eukaryotic genes, RNA polymerase II transcribes through the site where polyadenylation will occur. Consequently, the 3′ terminus of the mature mRNA is formed by site-specific posttranscriptional cleavage and polyadenylation (for review, see Birnstiel et al. 1985; Proudfoot and Whitelaw 1988; Proudfoot 1989). Although discrete sites for the termination of the primary transcript have not yet been identified, general regions of DNA a few hundred nucleotides in length and downstream from the polyadenylation site have been identified where transcription randomly terminates.

Two distinct sequence elements are required for accurate and efficient polyadenylation: (1) GU- or U-rich sequences located downstream from the polyadenylation site and (2) a highly conserved sequence of six nucleotides, AAUAAA, located 11–30 nucleotides upstream, which is necessary but not sufficient for posttranscriptional cleavage and polyadenylation (for review, see Mason et al. 1986; Proudfoot and Whitelaw 1988). The practical implication of these observations is that sequences downstream from the polyadenyl-

ation site must be included in eukaryotic expression vectors to ensure efficient polyadenylation of the mRNA of interest. Although a full-length cDNA clone may encode the conserved AAUAAA sequence and a tract of poly(A), these endogenous elements are not by themselv s sufficient to guarantee polyadenylation. The downstream GU- or U-rich sequences necessary for cleavage and polyadenylation must therefore be incorporated into the vector. The most frequently utilized signals are those derived from SV40; a 237-bp *Bam*HI-*Bcl*I restriction fragment contains the cleavage/polyadenylation signals from both the early and the late transcription units. These signals are positioned in opposite orientations, one on each DNA strand, and both sets of signals have been shown to be extremely efficient for the processing of hybrid mRNAs. Less frequently, polyadenylation signals have been provided by fusing a full-length cloned cDNA onto a partial genomic copy of a gene already resident in an expression vector (O'Hare et al. 1981; Kaufman et al. 1986b).

Sequences within the 3′ noncoding regions of eukaryotic genes may play a role in mRNA stability. For example, the presence of an AU-rich sequence, derived originally from the 3′ noncoding region of granulocyte-macrophage colony-stimulating factor (GM-CSF), has been shown to destabilize mRNAs transcribed from mammalian expression vectors (Shaw and Kamen 1986). Although similar motifs have been found in analogous locations within mRNAs encoding a variety of growth factors and oncogenes, relatively little is known about the way they function. To obtain maximal expression of a cloned gene, it may therefore be necessary to remove the nucleotide sequences 3′ of the termination codon.

SPLICING SIGNALS

The DNA sequences coding for a eukaryotic protein are rarely contiguous; usually, they are separated in the genome by intervening noncoding sequences that may vary in size from tens to many thousands of nucleotides. Following polyadenylation of the primary transcript, the introns are removed by splicing to generate the mature mRNA, which is then transported from the nucleus to the cytoplasm (for review, see Nevins 1983; Green 1986; Padgett et al. 1986; Krainer and Maniatis 1988).

The minimal sequences required for splicing of mRNA are located at the 5′ and 3′ boundaries of the intron. Comparison of a large number of these sequences has led to the identification of consensus sequences in which the first two and the last two nucleotides of the intron are essentially invariant:

$$\underset{\text{5′ splice site}}{\text{A G : G U (A) A G U}} \ldots \text{intron} \ldots \underset{\text{3′ splice site}}{\text{(U/C) } N_{11} \text{ C A G : G}}$$

The development of in vitro splicing systems has led to the elucidation of much of the biochemistry of the splicing reaction, but the processes that guarantee correct matching of 5′ and 3′ splice sites are not yet understood. The fact that hybrid pre-mRNAs containing 5′ and 3′ splice sites derived from different introns can be accurately spliced (Chu and Sharp 1981) indicates the importance of the conserved consensus sequences in this process. However, these sequences cannot be the sole determinants of splice-site selection, since identical, but ordinarily inactive, consensus sequences can be

found within both exons and introns of many eukaryotic genes. Such "cryptic" splice sites can be efficiently utilized when the normal splice sites are inactivated by mutation (Treisman et al. 1983; Wieringa et al. 1983).

Both the distance between splice sites and the DNA sequences surrounding them may influence the pathway of splicing in pre-mRNAs that contain multiple introns (Reed and Maniatis 1986). Alterations to the exon sequences flanking 5′ or 3′ splice sites can dramatically affect the efficiency with which the adjacent splice site is utilized. These findings are relevant to the design of eukaryotic expression vectors: Substitution of exon sequences or juxtaposition of normally noninteracting splice sites in a hybrid transcription unit might lead to the appearance of inappropriately spliced transcripts that cannot be translated.

Early studies of the expression of β-globin cDNA clones in cultured mammalian cells suggested that splicing is required for the production of cytoplasmic β-globin mRNA (Hamer and Leder 1979a,b,c). Furthermore, the expression of a gene with a mutation at a natural splice site could be rescued by insertion of a heterologous intron into the transcription unit (Gruss et al. 1979; Gruss and Khoury 1980). It is now known that this requirement for splicing signals is not absolute: Many cDNAs have been efficiently expressed from vectors that lack splicing signals (see, e.g., Gething and Sambrook 1981; Treisman et al. 1981). However, because the presence of an intron has proven to be deleterious in only a few cases and because some genes appear to be expressed more efficiently when introns are present, we recommend the use of vectors that contain a splice donor and acceptor site within the mammalian transcription unit.

ELEMENTS FOR REPLICATION AND SELECTION

In addition to the elements already described, eukaryotic vectors may contain other specialized elements intended to increase the level of expression of cloned genes or to facilitate the identification of cells that carry the transfected DNA.

Viral replicons

A number of animal viruses contain DNA sequences that promote the extrachromosomal replication of the viral genome in permissive cell types. Plasmids bearing these viral replicons are replicated episomally as long as the appropriate *trans*-acting factors are provided by genes either carried on the plasmid or within the genome of the host cell. Different viral replicons work with different efficiencies. Plasmid vectors containing the replicons of papovaviruses such as SV40 or polyomavirus replicate to extremely high copy number in cells that express the appropriate viral T antigen. Because the transfected cells die after 3 or 4 days, when the number of plasmid molecules exceeds 10^4 copies/cell, these systems are used for the transient, but abundant, expression of the transfected genes (see pages 16.17–16.22). Plasmid vectors containing replicons from viruses such as bovine papillomavirus (see pages 16.23–16.26) and Epstein-Barr virus (see pages 16.26–27) are propagated episomally at lower copy numbers (usually <100 copies/cell) and do not generally cause cell death. These vectors can be used to isolate stable

lines of cells that permanently express more modest levels of the transfected genes.

Genes encoding selectable markers

DNA, which enters only a small proportion of mammalian cells in a given culture, becomes stably maintained in an even smaller fraction. In a very few cases—for example, when the cells are transformed by an oncogene—stably transfected cells can be identified because they express an altered phenotype such as morphological transformation, loss of contact inhibition, or increased growth rate. However, in the great majority of cases, isolation of cell lines that express the transfected gene is achieved by introduction into the same cells of a second gene that encodes a selectable marker, i.e., an enzymatic activity that confers resistance to an antibiotic or other drug. Some of the markers described below are dominant and can be used with any type of mammalian cell; others must be used with particular cell lines that lack the relevant enzyme activity.

In early experiments, the genes encoding the protein of interest and the selectable marker were included on a single vector. However, Wigler et al. (1979) found that mammalian cells capable of taking up DNA do so efficiently, so that two unlinked plasmids can be cotransfected with high frequency ($>90\%$). Cotransfection, which obviates the need to construct complex recombinants, has become the standard method of introducing a selectable marker (on one plasmid) and the gene of interest (on another plasmid) into mammalian cells. The selectable markers that are currently used include:

- *Thymidine kinase.* The thymidine kinase gene (*tk*), which is expressed in most mammalian cells, codes for an enzyme that is involved in the salvage pathway for synthesis of thymidine nucleotides. A number of tk^- cell lines have been isolated from different mammalian species, including mouse (Ltk^- cells) (Kit et al. 1963; Wigler 1977), human ($143tk^-$ cells) (Bacchetti and Graham 1977), and rat (Rat-2 fibroblast cells) (Topp 1981). These mutant cell lines, in contrast to their wild-type parents, will grow in medium that contains the thymidine analog 5-bromodeoxyuridine. Szybalska and Szybalski (1962) and Littlefield (1964, 1966) developed a selective medium containing hypoxanthine, aminopterin, and thymidine (HAT medium; see Appendix A) in which only cells expressing the *tk* gene will grow. By the appropriate use of this medium, it is therefore possible to select for or against cells that express the *tk* gene.

  Early cotransfection experiments utilized purified fragments of herpes simplex virus (HSV) DNA that contained the viral *tk* gene (Wigler et al. 1977). Subsequent cloning of the *tk* gene both from HSV (Colbère-Garapin et al. 1979) and from chicken cells (Perucho et al. 1980) made it possible to construct plasmids such as that shown in Figure 16.1A for use in cotransfection experiments. The primary limitation of these vectors is that they can be used only in tk^- cell lines.

- *Dihydrofolate reductase.* Mutants of CHO cells that lack the enzym dihydrofolate reductase (Urlaub and Chasin 1980) cannot synthesize tetrahydrofolate and therefore can grow only in media supplemented with

thymidine, glycine, and purines. Transfection of these cells with vectors that express a cloned copy of the dihydrofolate reductase gene (*dhfr*) gives rise to clones that can grow in the absence of these supplements (Subramani et al. 1981; Kaufman and Sharp 1982a,b; Kaufman et al. 1985; see Figures 16.1B and 16.3C).

DHFR can be inhibited by methotrexate, a folate analog. Progressive selection of cells that are resistant to increasing concentrations of methotrexate leads to amplification of the *dhfr* gene, with concomitant amplification of extensive regions of the DNA that flank the *dhfr* sequences (Schimke 1982). DNAs that are cotransfected with the *dhfr* gene tend to become integrated into the same region of the cellular chromosome and therefore can frequently be coamplified with *dhfr*. Alternatively, cells lacking DHFR activity can be transfected with a recombinant construct containing the gene of interest linked to the *dhfr* gene. The linked gene is then amplified by selecting with successively higher concentrations of methotrexate. The resulting cell lines express very high levels of the desired recombinant protein product (Kaufman and Sharp 1982a,b; Kaufman et al. 1985). This approach is described in more detail on page 16.28.

The coamplification method has also been adapted for use with cells that synthesize wild-type levels of DHFR. In one approach, the *dhfr* gene was placed under the control of a strong promoter, thereby conferring on transfected cells the ability to grow in concentrations of methotrexate that would be lethal to cells expressing normal, wild-type levels of the enzyme (Murray et al. 1983). Alternatively, cells transfected with a plasmid that carries a dominant selectable marker (e.g., resistance to geneticin [G418]), the *dhfr* gene, and the gene of interest are selected first for their ability to grow in G418 and then for their ability to grow in progressively higher concentrations of methotrexate (Kim and Wold 1985). Finally, an altered form of the *dhfr* gene encoding an enzyme that is more resistant to methotrexate has been utilized as a dominant selectable marker for cotransformation experiments in a broad range of cell types (Spandidos and Siminovitch 1977; O'Hare et al. 1981; Simonsen and Levinson 1983).

Note: G418 is now commercially available. Because cultured lines of mammalian cells differ widely in their sensitivity to this antibiotic, the concentration appropriate for the selection of stably transfected cells must be determined empirically.

- *Aminoglycoside phosphotransferase.* The mostly widely used dominant selection system utilizes the bacterial gene encoding aminoglycoside 3′ phosphotransferase (APH). Two distinct APH enzymes, encoded by the bacterial transposons Tn*5* and Tn*601*, confer resistance to aminoglycoside antibiotics such as kanamycin, neomycin, and geneticin, which inhibit protein synthesis in both prokaryotic and eukaryotic cells. Eukaryotic cells do not normally express an endogenous APH activity, but they are capable of expressing the enzymes encoded by the bacterial transposons. When fused to eukaryotic transcriptional regulatory elements, the genes encoding APH can be used as dominant markers to select cells that take up exogenous DNA (Jimenez and Davies 1980; Colbère-Garapin et al. 1981). The first APH (*neo*^{r}) vectors designed for mammalian cells expressed the Tn*5* *neo*^{r} gene under the control of the HSV *tk* promoter and polyadenylation sequences (Colbère-Garapin et al. 1981). Subsequently, vectors were

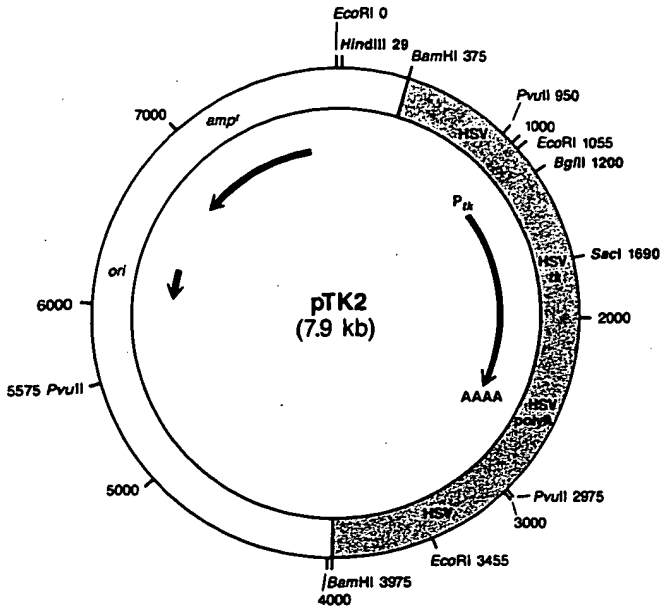


FIGURE 16.1A
pTK2 is a derivative of pBR322 that carries a 3.6-kb *Bam*HI fragment of herpes simplex virus (HSV) encoding thymidine kinase (*tk*). The positions of the *tk* promoter (P_{tk}) and the polyadenylation site (polyA; AAAA) are indicated.

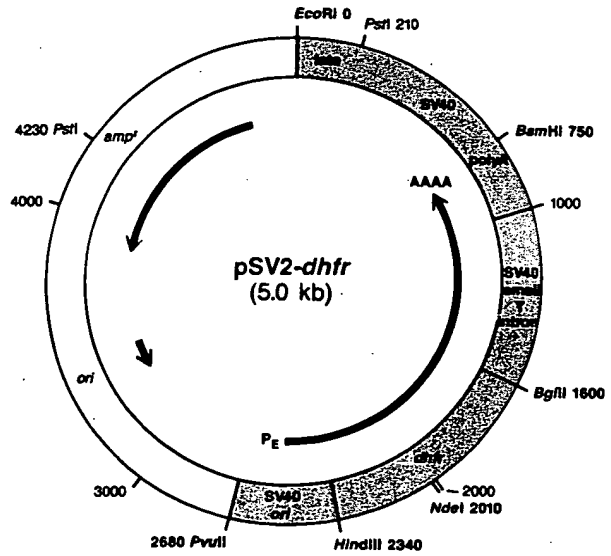


FIGURE 16.1B
pSV2-*dhfr* carries the SV40 origin (SV40 *ori*) and expresses dihydrofolate reductase (*dhfr*) from the SV40 early promoter (P_E). The SV40 small T intron and polyadenylation site (polyA; AAAA) are shown.

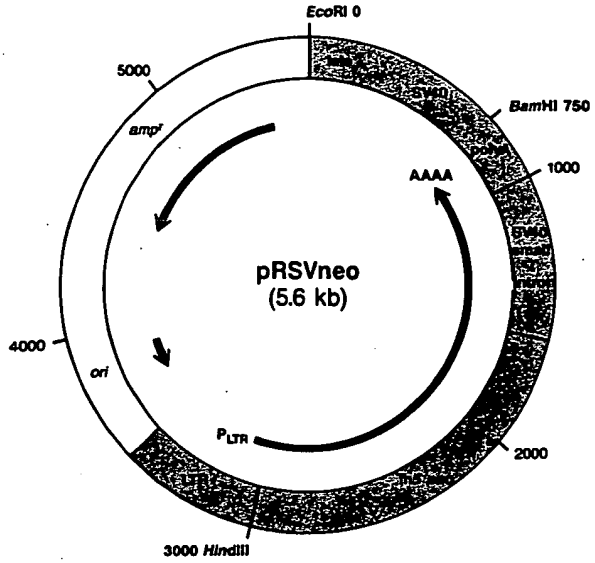


FIGURE 16.1C
pRSVneo expresses aminoglycoside phosphotransferase (APH) encoded by the bacterial transposon gene Tn*5* *neo*^{r} from the Rous sarcoma virus (RSV) LTR promoter (P_{LTR}). The SV40 small T intron and polyadenylation site (polyA; AAAA) are located downstream from Tn*5* *neo*.

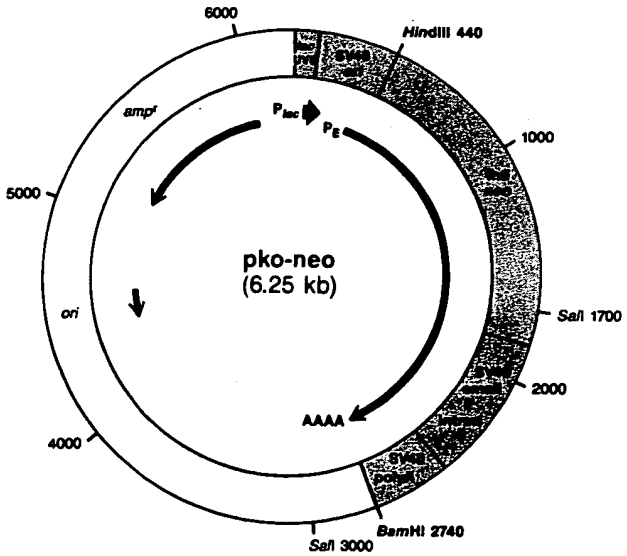


FIGURE 16.1D
pko-neo expresses aminoglycoside phosphotransferase encoded by the bacterial transposon gene Tn*5* *neo*^{r} from the eukaryotic SV40 early promoter (P_E) or the prokaryotic *E. coli lac*UV5 promoter (P_{lac}). The SV40 origin (SV40 *ori*), SV40 small T intron, and SV40 polyadenylation sites (polyA; AAAA) are present.

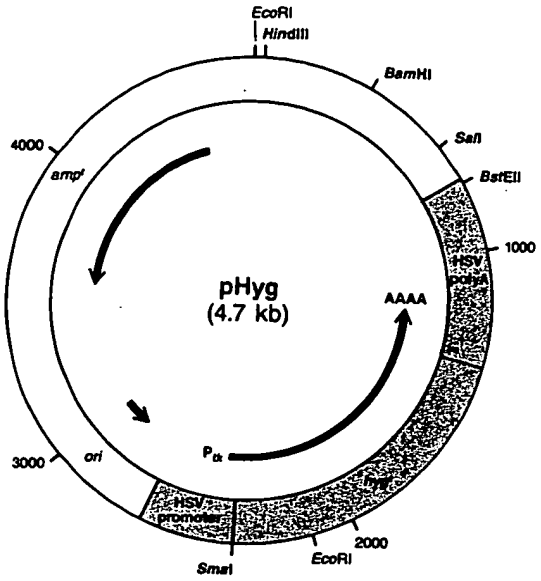


FIGURE 16.1E
pHyg directs the expression of the *E. coli* gene encoding hygromycin B phosphotransferase (*hyg*^{r}) using the herpes simplex virus promoter (P_{tk}) and polyadenylation site (HSV polyA; AAAA).

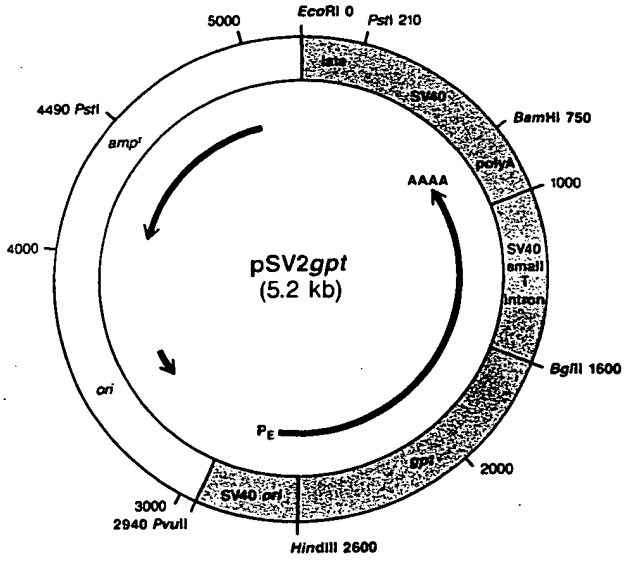


FIGURE 16.1F
In pSV2*gpt*, the *E. coli* xanthine-guanine phosphoribosyl transferase gene (*gpt*) is expressed using the SV40 early promoter (P_E) located in the SV40 origin (SV40 *ori*), the SV40 small T intron, and the SV40 polyadenylation site (polyA; AAAA).

developed that express the Tn*5* *neo*^{r} gene under the control of SV40 regulatory elements (Chia et al. 1982; Southern and Berg 1982; Okayama and Berg 1983; Van Doren et al. 1984). Vectors such as pSV2-*neo* (Southern and Berg 1982) and pRSVneo (Figure 16.1C), which have been widely used in cotransformation experiments, contain a version of the Tn*5* *neo*^{r} gene that retains prokaryotic promoter sequences between the eukaryotic promoter and the APH coding sequences. This configuration yields a vector that can confer antibiotic resistance upon both prokaryotic and eukaryotic cells. However, perhaps because the bacterial promoter contributes several upstream AUG codons, the efficiency of translation of APH mRNAs synthesized from these vectors is comparatively low in mammalian cells (Chen and Okayama 1987). Vectors such as pko-neo (Figure 16.1D) (Van Doren et al. 1984) and pcDneo (Okayama and Berg 1983; Chen and Okayama 1987), which lack prokaryotic promoter sequences, are therefore preferred.

- *Hygromycin B phosphotransferase.* The *E. coli* gene encoding hygromycin B phosphotransferase (Gritz and Davies 1983) can be used as a dominant selectable marker in much the same way as the APH gene. When the hygromycin B phosphotransferase gene (*hyg*) is introduced into mammalian cells on an appropriate expression vector (e.g., pHyg, Figure 16.1E) (Sugden et al. 1985), the transfected cells become resistant to the antibiotic hygromycin. Resistance to neomycin and to hygromycin can be selected for independently and simultaneously in cell lines that have been transfected with both genes. Thus, two different vectors can be introduced into one cell line, either simultaneously or sequentially.

- *Xanthine-guanine phosphoribosyl transferase.* The *gpt* gene of *E. coli* encodes the enzyme xanthine-guanine phosphoribosyl transferase (XGPRT), which is the bacterial analog of the mammalian enzyme hypoxanthine-guanine phosphoribosyl transferase (HGPRT). Whereas only hypoxanthine and guanine are substrates for HGPRT, XGPRT will also efficiently convert xanthine into XMP, which is a precursor of GMP. The bacterial *gpt* gene has been cloned and expressed in mammalian cells under the control of an SV40 promoter (Mulligan and Berg 1980, 1981a,b) (see, e.g., Figure 16.1F). Vectors expressing XGPRT restore the ability of mammalian cells lacking HGPRT activity to grow in HAT medium (Szybalska and Szybalski 1962; Littlefield 1964, 1966).

  Of much greater general use is the application of the *gpt* gene as a dominant selection system, which can be applied to any type of cell (Mulligan and Berg 1981a,b). Vectors expressing XGPRT confer upon wild-type mammalian cells the ability to grow in medium containing adenine, xanthine, and the inhibitor mycophenolic acid. Mycophenolic acid blocks the conversion of IMP into XMP and inhibits the de novo synthesis of GMP. The selection can be made more efficient by the addition of aminopterin, which blocks the endogenous pathway of purine biosynthesis.

- *CAD.* A single protein, CAD, possesses the first three enzymatic activities of de novo uridine biosynthesis (carbamyl phosphate synthetase, aspartate transcarbamylase, and dihydroorotase). Transfection of vectors expressing the CAD protein from Syrian hamsters into CAD-deficient (UrdA) mutants of CHO cells allows selection of CAD$^+$ transfectants that are able to grow in the absence of uridine (Robert de Saint Vincent et al. 1981).

L-Phosphonacetyl-L-aspartate (PALA) is a specific inhibitor of the aspartate transcarbamylase activity of CAD. Growth of wild-type or transfected mammalian cells in the presence of increasing concentrations of PALA leads to the amplification of the CAD gene and DNA sequences linked to it (Kempe et al. 1976; Robert de Saint Vincent et al. 1981; Wahl et al. 1984). The *E. coli* gene encoding aspartate transcarbamylase (*pyrB*), when expressed in CHO cells deficient in aspartate transcarbamylase, is also amplified by PALA selection (Ruiz and Wahl 1986).

- *Adenosine deaminase.* Adenosine deaminase (ADA) is present in virtually all animal cells, but it is normally synthesized in minute quantities and is not essential for cell growth. However, because ADA catalyzes the irreversible conversion of cytotoxic adenine nucleosides to their respective nontoxic inosine analogs, cells propagated in the presence of toxic concentrations of adenosine or its analog 9-β-D-xylofuranosyl adenine (Xyl-A) require ADA for survival (for references and review, see Kaufman 1987). Under conditions where ADA is required for cell growth, amplification of the gene can be achieved in the presence of increasing concentrations of 2′-deoxycoformycin (dCF), a transition-state analog of adenine nucleotides that strongly inhibits the enzyme. In cells selected for their ability to resist high concentrations of 2′-deoxycoformycin, it has been shown that ADA was overproduced 11,400-fold and represented 75% of the soluble protein synthesized by the cells (Ingolia et al. 1985).
- *Asparagine synthetase.* The *E. coli* gene coding for asparagine synthetase (AS) is a potentially useful, dominant, amplifiable marker for mammalian cells. Because the bacterial enzyme uses ammonia as an amide donor—in contrast to the mammalian enzyme, which uses glutamine—cells that express the bacterial AS gene will grow in asparagine-free medium containing the glutamine analog albizziin. Subsequently, the transfected AS gene can be amplified by selection in medium containing increasing concentrations of β-aspartyl hydroxamate, an analog of aspartic acid.

Foreign DNA Sequences

DNAs encoding the foreign protein of interest are usually cloned as cDNAs that lack all of the controlling elements required for expression in mammalian cells but may contain ancillary sequences introduced during the construction of the cDNA library (e.g., homopolymeric stretches of guanine or cytosine residues, synthetic linkers, etc.). No consensus exists as to whether or not these ancillary sequences need to be removed before the cDNA can be expressed in mammalian cells. However, since such sequences never enhance, and in some circumstances may suppress, the level of expression of foreign DNAs in mammalian cells (Simonsen et al. 1982), most workers prefer to remove as many extraneous sequences as is conveniently possible. Less frequently, DNAs encoding the foreign protein of interest are obtained as a genomic copy in which the coding sequences may be interrupted by one or more introns. A complete genomic copy will have all the controlling sequences necessary for the expression of the protein in some, but not necessarily all, cell types. Because the specificity of these sequences determines the range of cell types in which the gene will be active, replacement

of the enhancer and/or promoter sequences may be necessary to allow efficient expression of the gene in lines of cultured cells.

Cloned sequences that are expressed in mammalian cells almost invariably include the ribosome-binding site and initiation codon found in the natural gene. In contrast, in prokaryotic expression systems, the ribosome-binding site and initiation codon of the natural gene are almost invariably replaced by a Shine-Dalgarno ribosome-binding sequence and an ATG codon that have been optimally spaced for efficient expression. With mammalian mRNAs, the number of nucleotides between the starting sites of transcription and translation is not critical, but the initiation codon of the cloned gene should be the first AUG in the transcript. The presence of AUG codons upstream of the intended site of initiation may significantly decrease the efficiency of translation of the desired product (Hughes et al. 1984; Liu et al. 1984; Perez et al. 1987; Kozak 1989). In a transcription unit coding for two polypeptides, the efficiency of translation of the downstream reading frame is decreased several hundredfold, despite the presence of an in-frame termination codon between the two open reading frames (Kaufman et al. 1987). The consensus sequence for initiation of translation by eukaryotic ribosomes is the sequence

$$G\,C\,C\,G\,C\,C\,A^{-3}\,/\,G\,C\,C\,A^{1}\,U\,G\,G^{+4}$$

(Kozak 1989). However, Kozak points out that, for practical purposes, an initiaton codon can usually be designated "strong" or "weak" by considering only positions -3 and $+4$. As long as there is a purine in position -3, deviations from the rest of the consensus sequence only marginally impair initiation. In the absence of a purine in position -3, however, G^{+4} is essential for efficient translation and the contributions of other nearby nucleotides can be detected (Kozak 1989).

VECTOR SYSTEMS

A good way to establish rapidly the feasibility of expressing a cloned gene in mammalian cells is to use a transient expression system such as that provided by the simian COS cell line (Gluzman 1981). In addition to functional SV40 large T antigen, COS cells produce the permissivity factors required for replication of DNAs that contain the SV40 origin of DNA replication. In many cases, transient expression of cloned genes in COS cells will provide all of the experimental data required to solve the biological question at hand. In other situations, it may be necessary to choose another vector–host system—for example, when sustained high levels of production of a protein encoded by a cloned gene are required or when there is a need to obtain expression of a cloned gene in specialized cell types. Until recently, most mammalian expression vectors were designed to express a gene or cDNA of interest in a particular cell line. They were therefore tailored to contain controlling elements of the desired specificity and restriction sites to match those at the 5′ and 3′ termini of the sequences to be expressed. Consequently, most vectors in current use are idiosyncratic in their design and frequently do not contain either restriction sites or controlling elements that are universally useful. Construction of new recombinant vectors often requires extensive engineering (e.g., deletion or addition of restriction sites or substitution of controlling elements) of either the vector or the new gene.

In the following sections, various vector–host systems that are currently available are described in detail.

Plasmid-based Vectors That Do Not Carry a Eukaryotic Replicon

In these simple systems, a complete mammalian transcription unit (see pages 16.5–16.9) and a selectable marker (see pages 16.9–16.15) are inserted into a prokaryotic plasmid. The resulting vector is then amplified in bacteria before being transfected into cultured mammalian cells. Because these vectors do not contain a eukaryotic replicon, no episomal amplification of the transfected DNA occurs. Instead, the transfected DNA integrates into the genomes of a minority of the transfected cells, where it may direct the expression of low levels of the protein of interest. The selectable marker facilitates the isolation of the very small numbers of transfected cells that take up and express the foreign gene. Examples of simple vectors of this type that contain a range of different selectable markers and are suitable vehicles for the transfection of complete mammalian transcription units are shown in Figure 16.1 (e.g., pTK2, pHyg, and pRSVneo). In addition, all of the plasmids shown in Figure 16.1 can be used in cotransformation experiments to allow selection of those cells that have taken up the foreign DNA.

Plasmid DNA Expression Vectors Containing Regulatory Elements from Eukaryotic Viruses

SIMIAN VIRUS 40 VECTORS

SV40 is a member of the papova group of small, nonenveloped DNA viruses and causes lytic infection of permissive monkey cells. The virus has been studied extensively, and most aspects of its molecular biology are understood

in great detail (for review, see Tooze 1980). In cultured African green monkey kidney cells, SV40 undergoes a conventional replication cycle, in which the expression of the viral early genes is followed by replication of viral DNA, synthesis of the late viral proteins, and, finally, assembly of progeny virus particles. In rodent cells, infection is blocked at the level of DNA replication.

The viral genome is a covalently closed circular double-stranded DNA molecule of 5243 bp. The genome is functionally divided into early and late regions, which are transcribed from the two DNA strands in opposite directions. The early region is transcribed throughout the lytic cycle, and differential splicing generates two mRNAs that encode the large T and small T antigens. The late region is transcribed only after the onset of DNA replication and encodes the viral capsid proteins VP1, VP2, and VP3 from spliced mRNAs.

A number of plasmid-based expression vectors carry individual regulatory regions derived from SV40 but lack most of the coding region of the viral genome. After transfection into mammalian cells, foreign DNAs cloned into these vectors are transiently expressed, but no virus particles are produced. The regulatory region most commonly used is a 300-bp segment of SV40 DNA that lies between the viral early and late transcription units and contains a number of different controlling *cis* elements (see Figure 16.2) (for review, see McKnight and Tjian 1986). These elements include (1) the origin of DNA replication, (2) the promoters and sites of initiation of transcription of early and late mRNAs, (3) T-antigen binding sites involved in activation of the origin of replication and in autoregulation of early transcription, (4) a sequence of three G/C-rich, 21-bp direct repeats that are recognized by cellular transcription factors, and (5) two 72-bp direct repeats that comprise the SV40 early enhancer. In addition to these regulatory sequences, some SV40-based plasmid vectors contain a complete transcription unit that encodes the SV40 large T antigen, whose expression is required to activate the SV40 origin of replication in simian cells.

The basic properties of these vectors are:

- They have low to moderate levels of expression in a wide variety of transfected mammalian cells and high levels of expression in transfected COS cells
- Both genomic DNA and cDNA sequences can be expressed, and there are no constraints on the size of the foreign DNA that can be inserted
- They are generally used as a transient system, although cell lines that express low levels of the gene of interest can be isolated if a selectable marker is present in the expression vector or is cotransfected into the cells
- Typically, these vectors contain the SV40 origin of replication, a promoter with a broad host range (e.g., the SV40 early promoter), and a polyadenylation signal (almost always from SV40)
- Many vectors of this class also carry splice donor and acceptor signals (usually the intron from the SV40 small T antigen gene)

The early promoter and/or enhancer sequences work at low to moderate levels in a wide range of mammalian cells in the absence of DNA replication.

However, much higher levels of expression can be obtained in transfected simian cells, where the viral origin of DNA replication (minimum size 85 bp) drives replication of the plasmid when viral T antigen is supplied in *trans* (Gluzman 1981). The resulting increase in copy number of the plasmid DNA provides many more templates that become transcriptionally active over the course of the transfection. The use of origin-containing transient expression vectors was greatly facilitated by the development of lines of cells (COS cells) that were derived by transformation of simian CV-1 cells with an origin-defective SV40 genome (Gluzman 1981). These cells constitutively express wild-type SV40 T antigen and contain all of the cellular factors required to drive the replication of SV40 origin-containing plasmids. The efficiency of replication seems to be increased if the transfected plasmids lack certain ill-defined sequences that appear to interfere with replication in eukaryotic cells (Lusky and Botchan 1981). Most currently available vectors are free of these "poison" sequences. Over the course of a transfection experiment, COS cells accumulate $>10^5$ copies per cell of recombinant expression plasmids containing the SV40 origin of replication (Mellon et al. 1981) and express high levels of foreign DNA sequences. Expression in this system is transient because replication of the transfected plasmids continues unchecked until the cells die (at ~70–90 hours posttransfection), presumably because they cannot tolerate the high levels of extrachromosomally replicating DNA.

Transient expression in COS cells is the most widely used of all eukaryotic expression systems. The recombinant vectors are easy to construct and to use, and there are no constraints on the amount of DNA that can be inserted or on the use of genomic DNA sequences. This system has provided convenient positive verification of cDNA clones by expression (Toole et al. 1984; Wood et al. 1984), has facilitated rapid analysis of mutations introduced into cloned cDNAs (Mishina et al. 1984), and, following the pioneering efforts of Okayama and Berg (1982, 1983, 1985), has made possible the screening of cDNA libraries constructed in expression vectors to isolate cDNAs via expression of a desired activity (Lee et al. 1985; Wong et al. 1985; Aruffo and Seed 1987; Seed and Aruffo 1987). Expression screening is discussed on page 16.69.

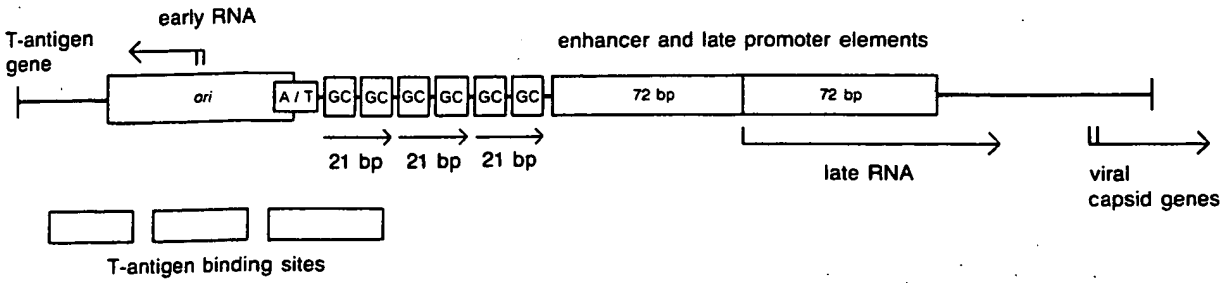


FIGURE 16.2
Regulatory elements in SV40-based expression vectors (see text for details). (Adapted from McKnight and Tjian 1986).

Although COS cells remain the most widely used host for transient expression of foreign DNAs, two more recently developed cell systems offer the possibility of controlling the replication of transfected plasmids. A line of temperature-sensitive COS cells produces high levels of thermolabile large T antigen under the transcriptional control of the Rous sarcoma virus LTR (Rio et al. 1985). These cells, which support replication of SV40-origin-containing vectors at 33°C but not at 40°C, can therefore be used to regulate the copy number of transfected plasmids. In addition, vectors that carry a selectable marker can be maintained either as integrated DNA or as autonomously replicating episomes. Similarly, Gerard and Gluzman (1985) have isolated transformed monkey cell lines (CMT and BMT) that express T antigen under the control of the murine metallothionein promoter. A direct correlation was found between the level of T antigen synthesis (which could be induced five- to tenfold by addition of heavy metals) and the extent of episomal replication of transfected plasmids containing the SV40 origin.

A number of vectors have been tailored to simplify expression of cDNAs in mammalian cells (Okayama and Berg 1982, 1983, 1985). In addition to the minimal sequences required for expression of cloned genes, many vectors used for transient expression in COS cells now carry ancillary sequences that simplify the cloning process or serve a variety of specialized functions. Included among these specialized sequences are: (1) polycloning sites; (2) selectable markers that can be used to establish permanent cell lines carrying the sequences of the recombinant plasmid; (3) bacteriophage promoters that can be used to generate transcripts of the foreign gene, which can then be translated in vitro; and (4) additional transcription units that encode the SV40 large T antigen, expression of which is required to activate the SV40 origin of replication in normal monkey cells. However, expression of an exogenous T antigen gene is unnecessary if COS cells are used as the hosts for transfection.

Figure 16.3 shows maps of three currently used transient expression vectors. pMSG (Figure 16.3A; available from Pharmacia) is suitable for expression of cDNA sequences inserted into a polycloning site downstream from the mouse mammary tumor virus LTR promoter, which can be induced with glucocorticoid in hormone-responsive cells. The plasmid contains the *E. coli gpt* gene, which is expressed from the SV40 early promoter. pSVT7 (Figure 16.3B) contains a polycloning site for insertion of cDNAs downstream from the SV40 early promoter. The vector also contains a promoter for bacteriophage T7 DNA-dependent RNA polymerase located upstream of the site of insertion of the foreign DNA sequence. Thus, mRNA can be generated in vitro (see Chapter 18, pages 18.81–18.85) for translation in cell-free systems.

An example of a more highly evolved vector, pMT2, which has been used both to obtain high levels of transient expression in COS cells and to establish lines of CHO cells that efficiently express foreign proteins (see page 16.29) is shown in Figure 16.3C (Kaufman et al. 1989). It contains eukaryotic regulatory elements from several different sources: (1) the SV40 origin and early gene enhancer, (2) the adenovirus major late promoter (Ad MLP) coupled to a cDNA copy of the adenovirus tripartite leader, (3) a hybrid intron consisting of a 5′ splice site from the first exon of the tripartite leader

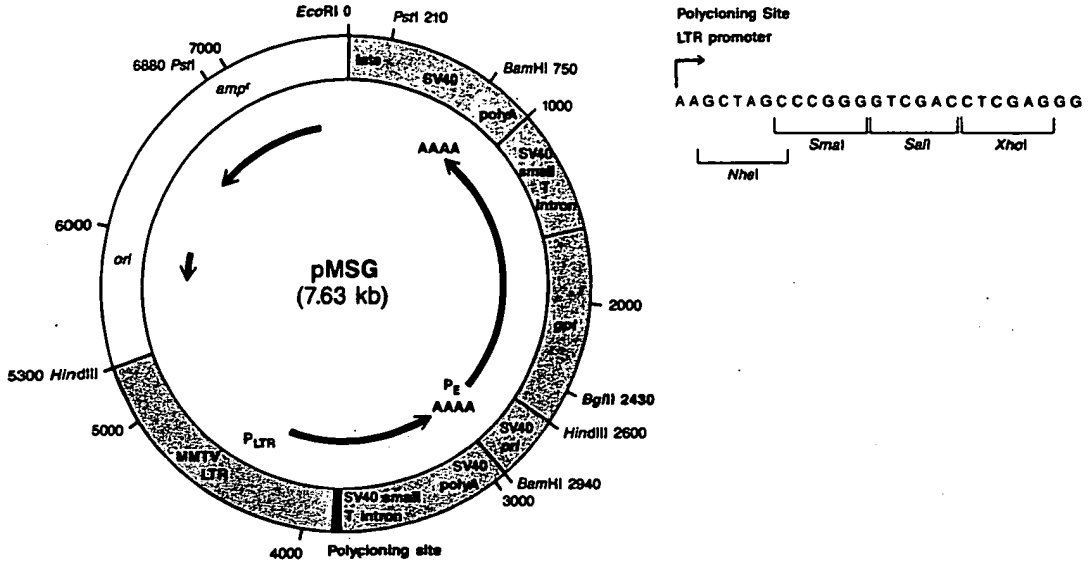


FIGURE 16.3A

pMSG is a transient expression vector that uses the mouse mammary tumor virus (MMTV) LTR promoter (P_{LTR}) to express cDNA sequences cloned into the polycloning site. The *E. coli* xanthine-guanine phosphoribosyl transferase gene (*gpt*) is expressed from the SV40 promoter (P_E) located in the SV40 origin (SV40 *ori*). The SV40 small T intron and polyadenylation site (SV40 polyA; AAAA) are found downstream from both the polycloning site and the *gpt* gene.

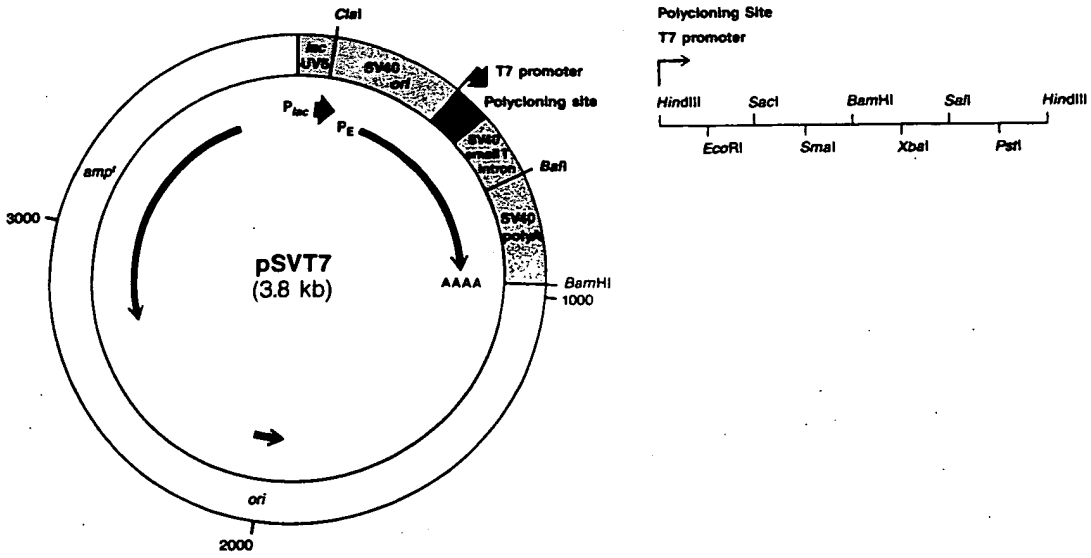


FIGURE 16.3B

pSVT7 is a transient expression vector that uses SV40 signals, the SV40 early promoter (P_E), the SV40 small T intron, and the SV40 polyadenylation site (polyA; AAAA) to express cDNA cloned into the polycloning site. The vector also carries the *lac*UV5 promoter (P_{lac}) for expression in *E. coli* and a bacteriophage T7 promoter for generation of mRNA in vitro.

and a 3′ splice site from a mouse immunoglobulin gene, (4) the SV40 polyadenylation signal, and (5) the adenovirus VA_I RNA gene region. pMT2 also carries sequences encoding murine DHFR positioned downstream from the splice acceptor site. cDNAs inserted between the splice acceptor site and the *dhfr* sequences can be transiently expressed at high levels in COS cells. The inserted cDNA is transcribed to produce a hybrid polycistronic mRNA in which the sequence coding for the foreign protein is flanked by the adenovirus tripartite leader and the murine *dhfr*. Because it lies at the 3′ terminus of the transcription unit, *dhfr* is inefficiently translated, but it can nevertheless serve as a selective, amplifiable marker (see page 16.28) and may also enhance the stability of the polycistronic mRNA (Kaufman et al. 1987 and unpubl.). The adenovirus tripartite leader and the VA_I RNA increase the efficiency of translation of the foreign coding sequences (3- to 20-fold) (Kaufman et al. 1985; Kaufman and Murtha 1987) by blocking the activity of a double-stranded RNA-dependent protein kinase that phosphorylates the α subunit of eukaryotic initiation factor 2 (eIF-2) (Kitajewski et al. 1986; O'Malley et al. 1986; Akusjärvi et al. 1987). (*Note:* Expression of the adenovirus VA RNAs is restricted to cells [e.g., COS cells] that contain a specific factor required for transcription of these RNA-polymerase-III-dependent genes.)

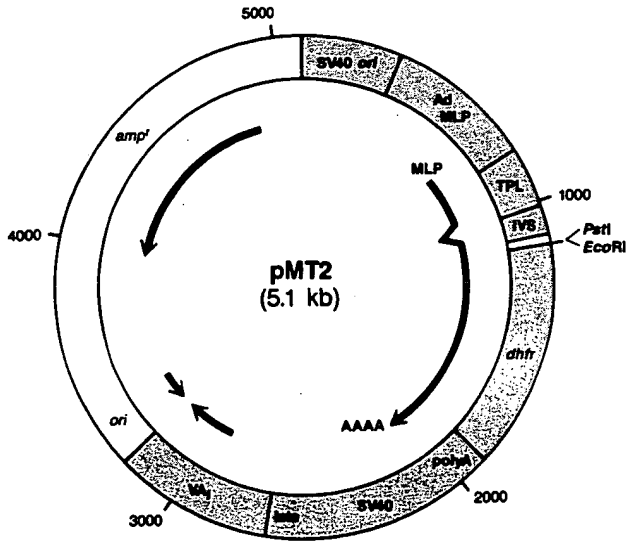


FIGURE 16.3C
pMT2 is a transient expression vector in which the adenovirus major late promoter (Ad MLP; MLP) is coupled to the adenovirus tripartite leader (TPL), which contains a 5′ splice site and part of a mouse immunoglobulin gene that contains a 3′ splice site. The intervening sequence (IVS) is followed by two cloning sites (*Pst*I and *Eco*RI), the dihydrofolate reductase gene (*dhfr*), the SV40 polyadenylation site (polyA; AAAA), and the adenovirus VA_I gene. The SV40 origin (SV40 *ori*) and enhancer are also present on the vector.

BOVINE PAPILLOMAVIRUS VECTORS

Papillomaviruses are small, double-stranded DNA viruses that commonly infect higher vertebrates, including man. These viruses usually cause benign, self-limiting, proliferative lesions commonly known as warts (for review, see Orth and Favre 1985), but they are oncogenic under certain circumstances (for review, see Broker and Botchan 1986). Many different papillomaviruses have been isolated, but most studies at the molecular level have been carried out with a bovine papillomavirus, BPV-1. This virus is capable of transforming rodent cells in vitro but, like all other papillomaviruses, will not grow in tissue culture. Molecular characterization of the BPV-1 genome was therefore achieved only after the viral DNA was cloned in prokaryotic vectors (Howley et al. 1980). The sequence of BPV-1 DNA was then rapidly determined (Chen et al. 1982), and the genomic organization of the virus was elucidated (Figure 16.4) (for review, see Broker and Botchan 1986).

When rodent cells in culture are morphologically transformed by BPV DNA, 20 to 100 copies of the viral DNA persist as extrachromosomal DNA. This ability of BPV DNA to replicate as an episome has led to its use both as an expression vector and as a vehicle for studying gene regulation. Not all of the viral genome is required for transformation or for establishment of the episomal state. The observation that both of these functions can be carried out by a fragment of the viral DNA (the 69% transforming fragment) led directly to the use of BPV DNA as a vector to express rat preproinsulin in murine C127 cells (Sarver et al. 1981). Since then, a number of BPV shuttle vectors have been developed that replicate episomally both in transformed mammalian cells and in bacteria (DiMaio et al. 1982; Sarver et al. 1982; for review, see Campo 1985).

The basic properties of these vectors are:

- They have low to moderate levels of expression in a wide variety of mammalian cells.
- Both genomic DNA and cDNA sequences can be expressed, and there are no constraints on the size of the foreign DNAs that can be inserted.
- They are never used as a transient expression system; instead, these vectors are used to establish cell lines that contain multiple copies of the foreign gene.
- Typically, these vectors contain a segment of BPV DNA (either the entire viral genome or the 69% transforming fragment), a promoter with a broad host range (e.g., the SV40 early promoter or the murine metallothionein promoter), a polyadenylation signal (almost always from SV40), splice donor and/or acceptor signals (usually the intron from the SV40 small T antigen gene), and "poisonless" plasmid sequences that allow the vector to be propagated in *E. coli*.
- BPV vectors may also carry a selectable marker (e.g., *neo*^{r}), whose expression is controlled by a separate promoter, and sequences derived from mammalian genomes which may help to maintain episomal replication of the recombinant vector in mammalian cells.

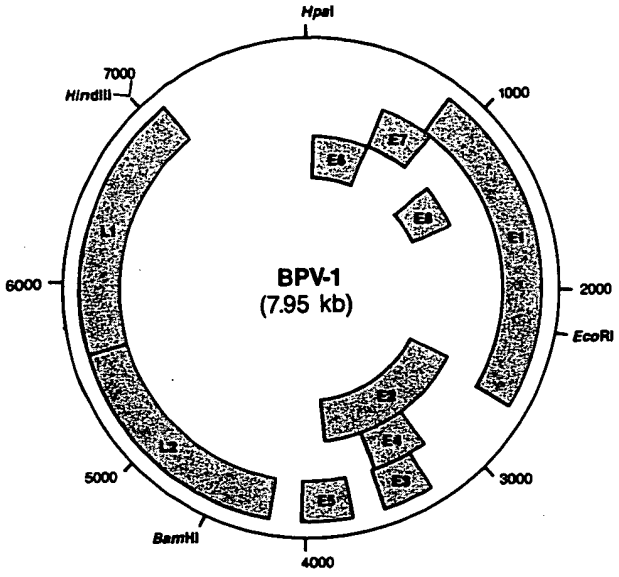


FIGURE 16.4
Bovine papillomavirus (BPV-1) encodes eight early gene products (E1–E8) and two late gene products (L1 and L2).

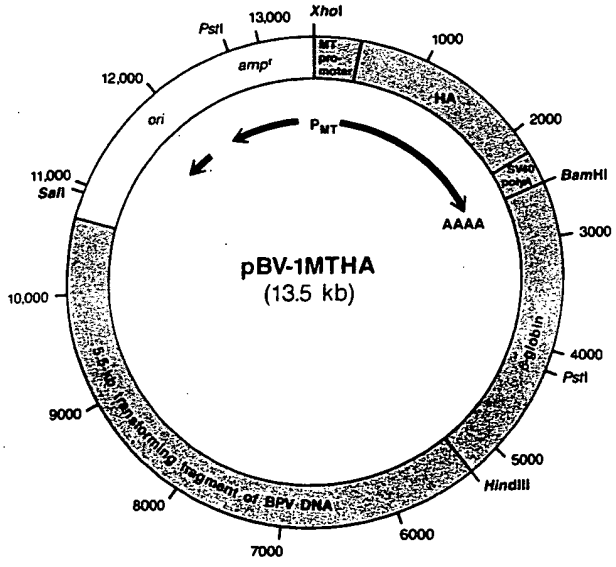


FIGURE 16.5
pBV-1MTHA carries the 69% transforming fragment (5.5-kb transforming fragment of BPV DNA), which allows episomal replication of this vector in mammalian cells. In addition, it encodes hemagglutinin (HA) expressed from the metallothionein (MT) promoter (P_{MT}) and human β-globin. It also has an SV40 polyadenylation site (SV40 polyA; AAAA).

Certain vectors (those of the BV-1 series [Figure 16.5]) carry an additional fragment of DNA derived initially from the human β-globin gene cluster. In some systems (DiMaio et al. 1982; Zinn et al. 1983; Sambrook et al. 1985), but not in all (DiMaio et al. 1984), this fragment appears to carry an activity that enhances the ability of BPV vectors containing only the 69% transforming fragment to replicate episomally in mammalian cells. Because of the large size of BPV-based vectors (8–14 kb) and the dearth of unique cloning sites, the insertion of a foreign gene into BPV vectors is not always straightforward. Most vectors contain only one or two restriction sites at which the coding sequences of interest may be inserted, attached either to their natural controlling elements or to heterologous controlling elements carried by the vector. To date, no BPV vectors have been produced that contain polycloning sites.

Following construction and amplification in bacteria, recombinant BPV DNAs are transfected into cultured mammalian cells, usually by the calcium phosphate coprecipitation technique (see pages 16.32–16.40). If a selectable marker is not used, cells that have been successfully transfected can sometimes be identified by their characteristic transformed morphology. Individual transformed clones are then isolated and tested for expression of the protein of interest. Typically, different transformants isolated from the same transfection vary up to tenfold in their level of expression of the protein of interest, and it is therefore worthwhile to screen a number of clones to identify those that express the protein at the desired level. By assaying a number of individual cell clones, it is usually possible to establish cell lines that yield 50- to 100-fold more protein than cell lines transformed with simple plasmid-based vectors (for review, see Campo 1985; DiMaio 1987).

In the absence of a selectable marker, the host range of BPV vectors is limited to those cells, such as the murine C127 line, that undergo identifiable changes in their morphology and growth after transfection by BPV DNA. Expansion of the host and tissue range to cells that are capable of supporting the autonomous replication of BPV-1-derived plasmids but do not manifest a transformed phenotype (e.g., mouse NIH-3T3, hamster CHO, canine MDCK, and pig PK-1 cell lines) can be achieved by the use of a dominant selectable marker. Several BPV vectors carry a copy of the bacterial neomycin resistance gene (*neo*^{r}) (see pages 16.10 and 16.14) under the control of efficient eukaryotic regulatory sequences (see, e.g., Law et al. 1983); alternatively, the *neo*^{r} gene may be cotransfected on a separate plasmid (e.g., pko-neo [Figure 16.1D]) (Sambrook et al. 1985). Resistance to high concentrations of toxic heavy metals such as cadmium is conferred by murine or human metallothionein genes and may also be used as a dominant selectable marker (Karin et al. 1983; Pavlakis and Hamer 1983; Krystal et al. 1986) (see Figure 16.5). Other potential markers, such as *tk*, *dhfr*, or *E. coli gpt*, have proven to be unsatisfactory because the BPV-based vector undergoes rearrangement and/or integrates into the cellular chromosome (for references, see DiMaio 1987).

The copy number and physical state of the recombinant genomes are affected both by the composition of the vector and by the host cell (Campo 1985; Sambrook et al. 1985) in ways that are not understood. Recombinant vectors containing the entire BPV genome or the 69% transforming fragment are usually propagated as stable, multicopy (20–300 copies/cell) extrachromosomal elements. Less frequently, the recombinant sequences are main-

tained as oligomeric plasmids or as head-to-tail tandem arrays integrated into the cellular chromosome. Whatever the state of the DNA, the cell lines invariably carry many copies of the gene of interest. It is this high copy number that is responsible, at least in part, for the efficient expression of the foreign protein.

EPSTEIN-BARR VIRUS VECTORS

The human herpesvirus, Epstein-Barr virus (EBV), transforms resting human B lymphocytes into dividing blast cells that can proliferate indefinitely in culture (for review, see Sugden 1982). EBV-transformed lymphoblasts secrete immunoglobulin, carry multiple episomal copies of the 172-kb viral DNA, and are usually diploid. The discovery that a *cis*-acting element of EBV, *oriP*, permits maintenance of episomal DNA molecules in adherent cells that carry EBV DNA (Sugden et al. 1985) and express the *trans*-acting EBNA-1 antigen (Lupton and Levine 1985; Yates et al. 1985) led to the development of vectors for the expression of foreign genes in a broad range of mammalian cells. The basic properties of these vectors are:

- They have low to moderate levels of expression in a wide variety of mammalian cells.
- Both genomic DNA and cDNA sequences can be expressed, and there are no constraints on the size of the foreign DNAs that can be inserted.
- They are never used as a transient expression system; instead, these vectors are used to establish cell lines that contain multiple episomal copies of the foreign gene.
- Typically, these vectors contain a segment of EBV DNA that carries the *oriP* region (a *cis*-acting replication element), plasmid sequences that allow the vector to be propagated in *E. coli*, restriction sites at which a foreign transcription unit may be inserted, and a selectable marker (e.g., hygromycin resistance) whose expression is controlled by a broad-host-range promoter.

One example of an EBV-based vector, pHEBo, is shown in Figure 16.6A. This vector has been used to express cloned human class I major histocompatibility complex genes in human lymphoblastoid cells (Shimizu et al. 1986). More complex vectors have been designed to extend the host range beyond EBV-transformed human cells (Yates et al. 1985). For example, p205 (Figure 16.6B) contains both *oriP* and a variant EBNA-1 gene that has sustained a spontaneous deletion of a repetitive sequence of glycine and alanine codons and supports plasmid replication more efficiently than the wild-type EBNA-1 gene. p205 replicates episomally in a variety of animal cells at a copy number of 1–90 molecules per cell, depending on the cell type used. In some rodent cells, however, the plasmid appears to integrate into the host-cell chromosome. The vector contains unique *Nar*I and *Nru*I sites that can be used for the insertion of a foreign transcription unit. A vector similar to p205 has been used in a shuttle cosmid vector to express and rescue the gene coding for human tumor necrosis factor (Kioussis et al. 1987).

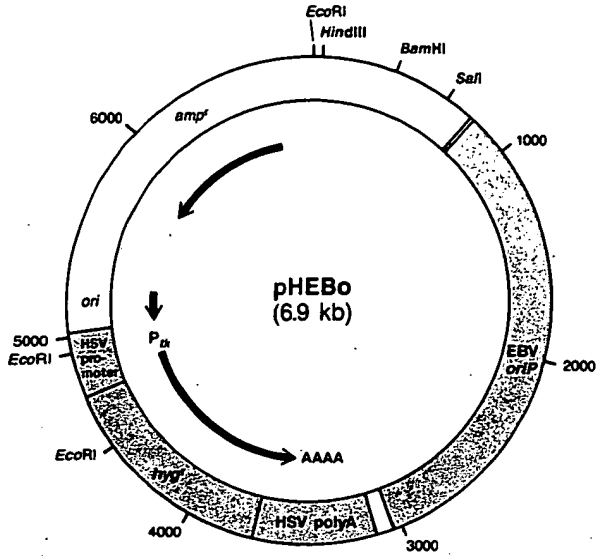


FIGURE 16.6A
pHEBo contains the Epstein-Barr virus origin *P* (EBV *oriP*) and directs the expression of the *E. coli* gene encoding hygromycin B phosphotransferase (hyg^r) using the herpes simplex virus promoter (P_{tk}) and polyadenylation site (HSV polyA; AAAA).

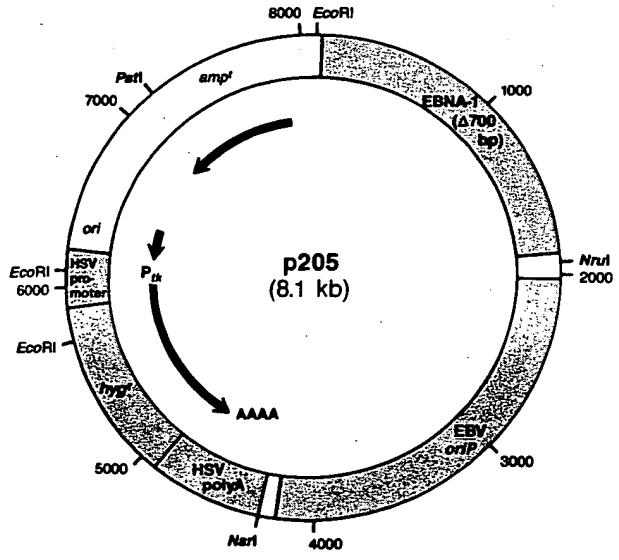


FIGURE 16.6B
p205 carries the Epstein-Barr virus origin *P* (EBV *oriP*) and expresses a *trans*-acting Epstein-Barr nuclear antigen (EBNA-1) from DNA containing a 700-bp deletion (Δ700 bp). It also directs the expression of the *E. coli* gene encoding hygromycin B phosphotransferase (hyg^r) using the herpes simplex virus promoter (P_{tk}) and polyadenylation site (HSV polyA; AAAA).

SV40-, BPV-, and EBV-based vectors can all be used to generate cell lines that contain multiple copies of a foreign gene and express moderate amounts of the protein of interest. Larger quantities of foreign protein can be obtained from lines of cells carrying amplified chromosomal copies of the gene of interest. In this method, the foreign gene is attached to a segment of DNA that carries a drug resistance marker and transfected into the cells, or the foreign gene and drug resistance marker are "cotransfected" (see page 16.9) into the cells. When cells carrying such composite structures are exposed to progressively increasing concentrations of the appropriate drug, sublines can be selected in which the number of copies of the drug resistance marker and the foreign gene are greatly amplified. Of the wide variety of available drug resistance markers (see pages 16.9–16.14; for a full listing, see Kaufman 1987), the *dhfr* gene is the most extensively used for coamplification of foreign DNA sequences. After several months of growth in progressively increasing concentrations of methotrexate, cell lines can be obtained that carry up to 1000 copies of the *dhfr* gene (for review, see Schimke 1984, 1988; Stark and Wahl 1984). The DNA unit that is amplified under selective conditions varies from cell line to cell line but is always much larger than the *dhfr* gene itself and may include up to 1000 kb of flanking DNA sequences. The amplified genes can be either stable or unstable when the cells are subsequently grown under nonselective conditions. In stable lines, the amplified genes are integrated into the chromosome and are associated with expanded chromosomal regions termed homogeneously staining regions (HSRs). In contrast, in unstably amplified lines, the *dhfr* genes are present on extrachromosomal, autonomously replicating elements called double-minute chromosomes (DMs), which do not contain centromeres and are rapidly lost upon propagation in the absence of selection. For reasons that are not well understood, hamster cell lines such as CHO generally contain stably amplified, integrated *dhfr* genes, whereas mouse cell lines generally carry the amplified sequences in the form of unstable DMs.

CHO cells that are deficient in DHFR (DUKX-B11 cells; Urlaub and Chasin 1980) can be transformed with a cloned *dhfr* gene and amplified as described above. This approach has been used to establish lines of CHO cells that express very high levels of proteins encoded by a number of cloned cDNAs (for review, see Kaufman 1987). To extend the *dhfr* amplification system to different cell types, a mutant *dhfr* gene that encodes a protein with a reduced sensitivity to methotrexate can be used to generate methotrexate-resistant derivatives of cell lines that contain normal numbers of the endogenous wild-type *dhfr* gene (see, e.g., Simonsen and Levinson 1983). However, the extent of amplification of this altered gene is limited by the high concentration of methotrexate required to inhibit the mutant *dhfr* gene. Thus, in practice, expression of foreign proteins from amplified gene copies has relied heavily upon the DHFR-deficient line of CHO cells. In principle, however, this method can be extended to any cell line by cotransfecting the *dhfr* gene and the foreign DNA together with a dominant selectable marker, for example, *neo*^{r} (Kim and Wold 1985). Cells that become resistant to neomycin are then exposed to progressively increasing concentrations of methotrexate as described above.

Foreign DNA sequences can be introduced into cells (1) as part of a simple *dhfr* vector such as pSV2-*dhfr* (see Figure 16.1B), (2) as part of a separate plasmid that is cotransfected with the *dhfr* vector, or (3) as part of a complex vector (e.g., pMT2) that can also be used to obtain transient expression in transfected COS cells. pMT2 (see Figure 16.3C) (Kaufman et al. 1989), which is described in detail on pages 16.20 and 16.22, contains the sequences coding for murine DHFR immediately downstream from the site at which foreign cDNAs are inserted. This arrangement results in the production of a dicistronic mRNA that carries the *dhfr* sequences towards its 3′ terminus. Originally, the *dhfr* sequences were placed in this position because they appeared to enhance the stability of the hybrid mRNA (Kaufman et al. 1985, 1986b; R. Kaufman, unpubl.). In these experiments, the separate plasmid pAdD26SV(A)-3 (Kaufman and Sharp 1982a,b), which contains a monocistronic *dhfr* transcription unit lacking an enhancer element, was cotransfected with a derivative of pMT2 [p91023(B)] containing the cDNA of interest. Under these conditions, it was believed that efficient expression of DHFR was dependent on continued physical association of the *dhfr* sequences with the enhancer provided by the p91023(B). Resistance to high concentrations of methotrexate would therefore require coamplification of the two separate transcription units. More recently, it has been shown that the *dhfr* gene can be translated from the distal region of the dicistronic transcript expressed from p91023(B), obviating the need for the second *dhfr* vector (Kaufman et al. 1987).

Two general procedures have been used to obtain cell lines carrying amplified copies of the foreign DNA sequences. First, individual clones of $DHFR^+$ transformants can be screened for expression of the heterologous gene and then amplified separately by growth in increasing concentrations of methotrexate. Alternatively, $DHFR^+$ transformants can be pooled, screened for expression of the gene of interest, and then grown en masse in the presence of progressively higher concentrations of methotrexate. In either case, the cells are then cloned, and the level of expression of the foreign protein in individual cell lines is measured. Although both of these procedures yield cell lines that produce large quantities of the protein of interest, Kaufman et al. (1985) found the second procedure to be better, perhaps because the mass selection favors clones that readily amplify the newly acquired DNA rather than clones that merely receive large amounts of DNA during transfection.

Although it has not yet been used frequently, an alternative amplification system based on selection for adenosine deaminase (ADA; see page 16.15) shows great promise (see, e.g., Yeung et al. 1985; Kaufman et al. 1986a). In contrast to *dhfr*, the gene encoding ADA is a dominant selectable marker and its use is therefore not limited to particular types of mammalian cells. An ADA coamplification vector has been constructed in which a transcription unit encoding ADA (Orkin et al. 1985) expressed from the SV40 early promoter has been introduced into a pMT2-based expression plasmid (Bonthron et al. 1986). The resultant expression vector is pMT3SV2.

Introduction of Recombinant Vectors into Mammalian Cells

Many methods have been developed for introducing cloned eukaryotic DNAs into cultured mammalian cells, several of which are discussed below.

- *Calcium phosphate- or DEAE-dextran-mediated transfection.* The most widely used method is transfection mediated by either calcium phosphate or DEAE-dextran. Although the mechanism remains obscure, it is believed that the transfected DNA enters the cytoplasm of the cell by endocytosis and is transferred to the nucleus. Depending on the cell type, up to 20% of a population of cultured cells can be transfected at any one time. Because of its high efficiency, transfection mediated by calcium phosphate or DEAE-dextran is the method of choice for experiments that require transient expression of the foreign DNA in large numbers of cells. Calcium phosphate-mediated transfection is also used to establish cell lines that carry integrated copies of the foreign DNA, which are usually arranged in head-to-tail tandem arrays.
- *Polybrene.* The polycation Polybrene allows the efficient and stable introduction of low-molecular-weight DNAs (e.g., plasmid DNAs) into cell lines (e.g., CHO cells) that are relatively resistant to transfection by other methods (Kawai and Nishizawa 1984; Chaney et al. 1986).
- *Protoplast fusion* (Schaffner 1980; Rassoulzadegan et al. 1982). In this method, protoplasts derived from bacteria carrying high numbers of copies of a plasmid of interest are mixed directly with cultured mammalian cells. After fusion of the cell membranes (usually with polyethylene glycol), the contents of the bacteria are delivered into the cytoplasm of the mammalian cells and the plasmid DNA is transferred to the nucleus. Protoplast fusion is not as efficient as transfection for many of the cell lines that are commonly used for transient expression assays, but it is useful for cell lines in which endocytosis of DNA occurs inefficiently. Protoplast fusion frequently yields multiple copies of the plasmid DNA tandemly integrated into the host chromosome (Robert de Saint Vincent et al. 1981).
- *Electroporation.* The application of brief, high-voltage electric pulses to a variety of mammalian and plant cells leads to the formation of nanometer-sized pores in the plasma membrane (Neumann et al. 1982; Zimmermann 1982). DNA is taken directly into the cell cytoplasm either through these pores or as a consequence of the redistribution of membrane components that accompanies closure of the pores. Electroporation can be extremely efficient and can be used both for transient expression of cloned genes and for establishment of cell lines that carry integrated copies of the gene of interest. Electroporation, in contrast to calcium phosphate-mediated transfection and protoplast fusion, frequently gives rise to cell lines that carry one, or at most a few, integrated copies of the foreign DNA (Boggs t al. 1986).

- *Liposomes.* Artificial membrane vesicles (liposomes) are being intensively studied for their usefulness as delivery vehicles in vitro and in vivo. For a review of the current procedures for liposome preparation, targeting, and delivery of contents, see Mannino and Gould-Fogerite (1988). Most of these procedures involve encapsulation of DNA or RNA within liposomes, followed by fusion of the liposomes with the cell membrane. However, Felgner et al. (1987) have reported that DNA that is coated with a synthetic cationic lipid can be introduced into cells by fusion. Although this method is simple and appears to be efficient, it is comparatively new and untested (but see Felgner and Holm 1989; Maurer 1989).
- *Direct microinjection into nuclei* (Capecchi 1980). Although this method has the advantage of not exposing DNA to cellular compartments such as low-pH endosomes, it cannot be used to introduce DNA on a scale large enough for biochemical analysis. Microinjection is therefore used primarily as a method to establish lines of cells that carry integrated copies of the DNA of interest.

Irrespective of the method used to introduce DNA into cells, the efficiency of transient or stable transfection is determined largely by the cell type that is used. Different lines of cultured cells vary by several orders of magnitude in their ability to take up and express exogenously added DNA. Furthermore, a method that works well for one type of cultured cell may be useless for another. Many of the techniques described below have been optimized for the standard lines of cultured cells. When using more exotic lines of cells, it is important to compare the efficiencies of several different methods. Below, we present detailed protocols for the most commonly used transfection techniques (calcium phosphate coprecipitation and DEAE-dextran transfection) and we outline alternative methods (transfection using Polybrene [page 16.47], protoplast fusion [pages 16.48–16.53], and electroporation [16.54–16.55]) that can be used with cell lines that are resistant to transfection by more standard techniques.

TRANSFECTION OF COPRECIPITATES OF CALCIUM PHOSPHATE AND DNA

The uptake of DNA by cells in culture is markedly enhanced when the nucleic acid is presented as a calcium phosphate–DNA coprecipitate. Graham and van der Eb (1973), who developed the procedure for the introduction of adenovirus and SV40 DNA into adherent cells, described the concentrations of calcium (125 mM) and DNA (5–30 μg/ml) that were optimal for formation of calcium phosphate–DNA coprecipitates at neutral pH (7.05). In addition, they established the optimal times for the precipitation reaction (20–30 minutes) and for the subsequent exposure of cells to the precipitate (5–24 hours). Their work laid the foundation for the introduction of cloned DNA into many different kinds of mammalian cells and led directly to reliable methods for stable transformation of cells and for transient expression of cloned DNAs. Many minor modifications of the procedure have been described, mostly involving permutations of the order and manner of mixing of ingredients in the precipitation reaction. Increases in the efficiency of the procedure have been achieved by incorporating additional steps such as glycerol shock (Parker and Stark 1979) and/or chloroquine treatment (Luthman and Magnusson 1983) following the transfection protocol. Treatment with sodium butyrate has also been shown to enhance the expression in simian and human cells of plasmids that contain the SV40 enhancer (Gorman et al. 1983a,b).

We describe below variants of calcium phosphate-mediated transfection that are used to introduce DNA into (1) adherent cells, (2) adherent cells that have been released from the substratum with trypsin, and (3) nonadherent cells. The second procedure, although applicable for all types of adherent cells, appears to be of particular advantage for polarized epithelial cells, which do not efficiently take up material by endocytosis through the apical plasma membrane, which is normally exposed to the culture medium. The final protocol differs from the standard protocol in that the calcium phosphate–DNA coprecipitate is formed slowly in the tissue culture medium during prolonged incubation (15–24 hours) under controlled conditions of pH (6.96) and CO_2 tension (2–4%). This is a highly efficient method (Chen and Okayama 1987, 1988) to achieve stable transformation of mammalian cells with supercoiled plasmid DNAs.

Standard Protocol for Calcium Phosphate-mediated Transfection of Adherent Cells

1. Prepare the following solutions:

2× HEPES-buffered saline (HBS)

280 mM NaCl
10 mM KCl
1.5 mM $Na_2HPO_4 \cdot 2H_2O$
12 mM dextrose
50 mM HEPES

Dissolve 1.6 g of NaCl, 0.074 g of KCl, 0.027 g of $Na_2HPO_4 \cdot 2H_2O$, 0.2 g of dextrose, and 1 g of HEPES in a total volume of 90 ml of distilled H_2O. Adjust the pH to 7.05 with 0.5 N NaOH, and then adjust the volume to 100 ml with distilled H_2O. Sterilize the solution by passage through a 0.22-micron filter. Store in 5-ml aliquots at −20°C.

2 M $CaCl_2$

Dissolve 10.8 g of $CaCl_2 \cdot 6H_2O$ in 20 ml of distilled H_2O. Sterilize the solution by passage through a 0.22-micron filter. Store in 1-ml aliquots at −20°C.

0.1× TE (pH 8.0)

1 mM Tris·Cl (pH 8.0)
0.1 mM EDTA (pH 8.0)

Sterilize the solution by passage through a 0.22-micron filter. Store in aliquots at 4°C.

DNA

Dissolve the DNA (~20 $\mu g/10^6$ cells) in 0.1× TE (pH 8.0) at a concentration of 40 $\mu g/ml$.

To obtain the highest transformation efficiencies, plasmid DNAs should be purified by equilibrium centrifugation in CsCl–ethidium bromide density gradients. If smaller amounts of DNA are used, carrier DNA should be added to adjust the concentration to 40 $\mu g/ml$. Eukaryotic carrier DNA prepared in the laboratory (see Chapter 9) usually gives higher transfection efficiencies than commercially available DNA such as calf thymus or salmon sperm DNA. Carrier DNA should be sterilized before use by ethanol precipitation or extraction with chloroform.

2. Twenty-four hours before transfection, harvest exponentially growing cells by trypsinization and replate them at a density of 1×10^5–2×10^5 cells/cm^2 in 60-mm tissue culture dishes in the appropriate serum-containing medium. Incubate the cultures for 20–24 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

3. For each 60-mm monolayer of cells to be transfected, prepare the calcium phosphate–DNA coprecipitate as follows: Place 220 μl of the DNA prepared in step 1 (40 μg/ml in 0.1 × TE [pH 8.0]) and 250 μl of 2 × HBS in a disposable, sterile 5-ml plastic tube (e.g., Falcon 2509). Slowly add 31 μl of 2 M $CaCl_2$ (with gentle mixing over a period of ~30 seconds).

   Incubate the mixture for 20–30 minutes at room temperature, during which time a fine precipitate should form. At the end of the incubation, pipette the mixture up and down once to resuspend the precipitate.

   A commonly used alternative procedure is to add a dilute mixture of $CaCl_2$ and DNA to the 2 × HBS solution. For example, to 250 μl of 2 × HBS, add dropwise with gentle mixing 250 μl of the DNA prepared in step 1 and $CaCl_2$ (250 mM). Incubate the mixture as described above.

   Both types of reaction mixtures can be doubled or quadrupled in volume if a larger number of cells are to be transfected. When the volumes are quadrupled, a larger tube (e.g., a 15-ml Falcon 2059 tube) should be used. Normally, 0.5 ml of the calcium phosphate–DNA coprecipitate is added to 5 ml of medium in a 25-cm^2 flask or 60-mm dish; 1 ml of the precipitate is normally added to 10 ml of medium in a 90-mm dish.

   Published procedures differ widely in the manner and rate of mixing of ingredients. Some advise against anything but the gentlest agitation and suggest that air, bubbled from an electric pipetting device, should be used to mix the solution. Others advocate continuous slow mixing during addition of the DNA solution, followed by gentle vortexing. The object is to avoid the rapid formation of coarse precipitates that results in a decreased efficiency of transformation. In practice, several factors other than the speed of mixing affect the size of the precipitate, including the concentration and size of the DNA (high-molecular-weight DNA can be sheared by passage through a fine syringe needle) and the exact pH of the buffer (some workers make up several batches of HBS buffer over the pH range 6.95–7.1 and test each batch for the quality of precipitates and for the efficiency of transfection). If it is crucial to achieve the highest transfection efficiencies, time should be spent to optimize these factors for your particular system. Once an efficacious batch of reagents has been prepared and stored as indicated above, reproducible results can be obtained over long periods of time.

4. Transfer the calcium phosphate-DNA suspension into the medium above the cell monolayer. (Use 0.5 ml of suspension for 5 ml of medium in a 60-mm dish [see note 2 to step 3].) Rock the dish gently to mix the medium, which will become yellow-orange and turbid. An alternative method is to remove the medium and add the precipitate directly to the cells. Incubate the cells for 15 minutes at room temperature, and then add medium to the dish. Following either procedure, incubate the transfected cells for up to 24 hours (depending on the choice of subsequent treatments; see step 5) at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

5. The transfected cells can then be treated in one of the following ways:

   a. If no additional treatments (with reagents such as chloroquine, glycerol, or sodium butyrate; see below) are to be employed, incubate the cells for 16–24 hours; then remove the medium and precipitate by aspiration. Wash the monolayer once with phosphate-buffered saline (PBS; see Appendix B), and add 5 ml of prewarmed complete growth medium. Return the cells to the incubator for 24–60 hours before

assaying for transient expression of the transfected DNA or replating the cells in the appropriate selective medium for the isolation of stable transformants.

b. In many instances, uptake of DNA is increased by concurrent treatment of the cells with chloroquine, which may act by inhibiting the degradation of the DNA by lysosomal hydrolases (Luthman and Magnusson 1983). The concentration of chloroquine to be added to the growth medium and the time of treatment are limited by the sensitivity of the cells to the toxic effect of the drug. The optimal concentration of chloroquine for the particular cell type that is used should therefore be determined in preliminary experiments. However, most cell types respond well to treatment with chloroquine at a final concentration of 100 μM for 3–5 hours. A stock solution (100 mM) of chloroquine diphosphate is diluted (1:1000) directly into the medium either before or after the addition of the calcium phosphate–DNA coprecipitate to the cells (see alternative procedures in step 4). During treatment with chloroquine, it is normal for cells to develop a vesicularized appearance. After the 3–5-hour treatment with DNA and chloroquine, remove the medium and the precipitate, wash with PBS, and add 5 ml of prewarmed complete growth medium. Return the cells to the incubator for 24–60 hours before assaying for transient expression of the transfected DNA or replating the cells in appropriate selective medium for the isolation of stable transformants.

   The 100 mM stock solution of chloroquine diphosphate (60 mg/ml in water) should be sterilized by filtration and stored in foil-wrapped tubes at −20°C).

c. Brief treatment of transfected cells with glycerol has also been shown to increase the efficiency of transformation or transient expression of the introduced DNA. This procedure may be used following treatment with chloroquine. Because cells vary widely in their sensitivity to the toxic effects of glycerol, each cell type must be tested in advance to determine the optimum time (30 seconds to 3 minutes) of treatment. Tolerant cells may be given a glycerol shock after 3–5 hours of exposure of the cells to the calcium phosphate–DNA coprecipitate in growth medium (+/− chloroquine).
   i. Remove the growth medium by aspiration and wash the monolayer once with PBS.
   ii. Add 1.5 ml of 15% glycerol in 1× HBS to the monolayer, and incubate the cells for 30 seconds to 3 minutes at 37°C.
   iii. Remove the glycerol by aspiration, and wash the monolayers once with PBS.
   iv. Add 5 ml of complete growth medium, and incubate the cells for 24–60 hours before assaying for transient expression of the transfected DNA or replating the cells in the appropriate selective medium for the isolation of stable transformants.

d. Sodium butyrate has also been shown to enhance the expression of recombinant plasmids carrying the SV40 early promoter/enhancer in simian and human cells (Gorman et al. 1983a). Cells are exposed to sodium butyrate following glycerol shock by adding the agent directly

to the growth medium (step 5c, iv). Different concentrations of sodium butyrate (prepared as a 500 mM stock solution by neutralizing butyric acid in a chemical hood with NaOH) are used, depending on the cell type, for example:

CV-1	10 mM
NIH-3T3	7 mM
HeLa S3	5 mM
CHO	2 mM

Incubate the cells for 24–60 hours before assaying for transient expression or replating the cells in the appropriate selective medium for the isolation of stable transformants.

6. a. *Transient expression:* Harvest the cells 48–60 hours after transfection for analysis of RNA or DNA by hybridization. Newly synthesized protein may be analyzed by radioimmunoassay, by western blotting, by immunoprecipitation following in vivo metabolic labeling, or by assays of enzymatic activity in cell extracts. For assays that involve replicate samples or treatment of transfected cells under multiple conditions or over a time course, it is desirable to avoid dish-to-dish variation in transfection efficiency. In these cases, it is best to transfect large monolayers of cells (90-mm dishes) and then to trypsinize the cells after 24 hours of incubation and distribute them among several smaller dishes.

b. *Stable transformation:* Following 18–24 hours of incubation in nonselective medium to allow expression of the transferred gene(s) to occur, the cells are trypsinized and replated in the appropriate selective medium. This medium should be changed every 2–4 days for 2–3 weeks to remove the debris of dead cells and to allow colonies of resistant cells to grow.

Individual colonies may be cloned and propagated for assay (for methods, see Jakoby and Pastan 1979). A permanent record of the numbers of colonies may be obtained by fixing the remaining cells with ice-cold methanol for 15 minutes and then staining them with 10% Giemsa for 15 minutes at room temperature before rinsing in tap water. The Giemsa stain should be freshly prepared in PBS or water and filtered through Whatman No. 1 filter paper.

The dilution at which the cells should be replated to yield well-separated colonies will be determined by the efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection.

Calcium Phosphate-mediated Transfection of Adherent Cells in Suspension

1. Prepare the solutions and form the calcium phosphate–DNA coprecipitate as described on pages 16.33–16.34, steps 1 and 3.

2. While the precipitate is forming, harvest exponentially growing adherent cells by trypsinization. Resuspend the cells in medium containing serum, and centrifuge aliquots that contain 10^6 cells at 800*g* for 5 minutes at 4°C. Discard the supernatant.

3. Resuspend each aliquot of 10^6 cells in 0.5 ml of the calcium phosphate–DNA suspension, and incubate for 15 minutes at room temperature.

4. To each aliquot, add 4.5 ml of prewarmed complete medium (with or without chloroquine; see page 16.35), and plate the entire suspension (~5 ml) in a single 90-mm tissue culture dish. Incubate the cells for up to 24 hours (depending on the choice of subsequent treatments; see step 5, pages 16.34–16.35) at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

Note

This technique can easily be modified to accommodate greater numbers of cells. For example, Chu and Sharp (1981) used 10^8 cells in 2 ml of calcium phosphate–DNA suspension containing 25 μg of DNA. After 15 minutes, the mixture was diluted with 40 ml of complete medium supplemented with 0.05 × HBS and 6.25 mM $CaCl_2$, and the cells were plated at 5×10^7 cells per 150-mm dish.

1. Prepare the solutions and form the calcium phosphate–DNA coprecipitate as described on pages 16.33–16.34, steps 1 and 3.

2. While the precipitate is forming, collect the exponentially growing cells by centrifugation at 800*g* for 5 minutes at 4°C. Discard the supernatant, and resuspend the cell pellet in 20 volumes of ice-cold phosphate-buffered saline (PBS; see Appendix B). Divide the suspension into aliquots that contain 1×10^7 cells. Recover the washed cells by centrifugation, and again discard the supernatant.

3. Gently resuspend 1×10^7 cells in 1 ml of calcium phosphate–DNA suspension (containing ~20 μg of DNA), and let the suspension stand for 20 minutes at room temperature.

4. Add 10 ml of serum-containing medium (with or without chloroquine; see page 16.35), and plate the entire suspension in a single 90-mm tissue culture dish. Incubate the cells for 6–24 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

5. Recover the cells by centrifugation at 800*g* for 5 minutes at room temperature, and wash them once with PBS. Resuspend the cells in 10 ml of prewarmed complete growth medium. Return the cells to the incubator for 48 hours before assaying for transient expression of transfected genes (step 6a, page 16.36) or replating the cells in selective medium for isolation of stable transformants (step 6b, page 16.36).

Note

If the cells have been proved to survive treatment with glycerol, a glycerol shock (detailed below) may be used 4–6 hours after exposure of the cells to the calcium phosphate–DNA coprecipitate begins in order to improve the frequency of transfection:

1. Collect the cells by centrifugation as described in step 5, and wash them once with PBS.
2. Resuspend the washed cells in 1 ml of 15% glycerol in 1× HBS, and incubate for 30 seconds to 3 minutes at 37°C.
3. Dilute the suspension with 10 ml of PBS, and recover the cells by centrifugation as described in step 5. Wash the cells in PBS.
4. Resuspend the cells in 10 ml of serum-containing medium, and plate them in a 90-mm tissue culture dish. Incubate the culture for 48 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2 before assaying for transient expression of transfected genes or replating the cells in selective medium.

Modified Calcium Phosphate-mediated Transfection Procedure

This is a highly efficient method to obtain stable transformation of mammalian cells with supercoiled plasmid DNAs. The calcium phosphate–DNA coprecipitate is allowed to form in the tissue culture medium during prolonged incubation (15–24 hours) under controlled conditions of pH (6.96) and CO_2 tension (2–4%) (Chen and Okayama 1987, 1988). It should be noted that linear DNAs yield very low transformation frequencies using this protocol, perhaps because the slow formation of the calcium phosphate–DNA coprecipitate delays protection of the DNA from nucleases. The best results have been obtained using supercoiled plasmid DNAs purified by two rounds of equilibrium centrifugation in CsCl–ethidium bromide density gradients.

1. Prepare the following solutions:

2 × BES-buffered saline

50 mM BES (*N*,*N*-bis[2-hydroxyethyl]-2-aminoethanesulfonic acid)
280 mM NaCl
1.5 mM $Na_2HPO_4 \cdot 2H_2O$

Dissolve 1.07 g of BES, 1.6 g of NaCl, and 0.027 g of Na_2HPO_4 in a total volume of 90 ml of distilled H_2O. Adjust the pH of the solution to 6.96 with HCl at room temperature, and then adjust the volume to 100 ml with distilled H_2O. Sterilize the solution by passage through a 0.22-micron filter, and store in aliquots at −20°C.

2.5 M $CaCl_2$

Dissolve 13.5 g of $CaCl_2 \cdot 6H_2O$ in 20 ml of distilled H_2O. Sterilize the solution by passage through a 0.22-micron filter. Store in 1-ml aliquots at −20°C.

2. Twenty-four hours before transfection, harvest exponentially growing cells by trypsinization and replate aliquots of 5×10^5 cells onto 90-mm tissue culture dishes. Add 10 ml of complete growth medium, and incubate the cultures overnight at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

3. Mix 20–30 μg of superhelical plasmid DNA with 0.5 ml of 0.25 M $CaCl_2$, add 0.5 ml of 2 × BES-buffered saline, and incubate the mixture for 10–20 minutes at room temperature. Do not expect a visible precipitate to form during this time.

4. Add the $CaCl_2$/DNA/BES-buffered saline solution dropwise to the dishes of cells, swirling gently to mix well. Incubate the cultures for 15–24 hours at 37°C in a humidified incubator in an atmosphere of 2–4% CO_2. The calcium phosphate–DNA complex forms slowly in the medium under conditions of low pH and precipitates gradually onto the cells during the incubation in an atmosphere containing low concentrations of CO_2. The nature of the precipitate is affected by the amount of DNA used. A

transition from a coarse precipitate to a fine precipitate occurs at the optimal DNA concentration (2–3 μg/ml in the growth medium).

5. Remove the medium by aspiration, and rinse the cells twice with medium. Add 10 ml of fresh medium, and incubate the cultures for 24 hours at 37°C in a humidified incubator in an atmosphere of 5% CO_2.

6. Following 18–24 hours of incubation in nonselective medium to allow expression of the transferred gene(s) to occur, the cells are trypsinized and replated in the appropriate selective medium. This medium should be changed every 2–4 days for 2–3 weeks to remove the debris of dead cells and to allow colonies of resistant cells to grow.

   The dilution at which the cells should be replated to yield well-separated colonies will be determined by the efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection.

7. Individual colonies may be cloned and propagated for assay (for methods, see Jakoby and Pastan 1979). A permanent record of the numbers of colonies may be obtained by fixing the remaining cells with ice-cold methanol for 15 minutes and then staining them with 10% Giemsa for 15 minutes at room temperature before rinsing in tap water. The Giemsa stain should be freshly prepared in PBS or water and filtered through Whatman No. 1 filter paper.

TRANSFECTION MEDIATED BY DEAE-DEXTRAN

DEAE-dextran was originally used as a facilitator to introduce poliovirus RNA (Vaheri and Pagano 1965) and SV40 and polyomavirus DNAs (McCutchan and Pagano 1968; Warden and Thorne 1968) into cells. The procedure, with slight modifications (see below), continues to be widely used for transfection of viral genomes and plasmids carrying viral sequences. Although the mechanism of action of DEAE-dextran is not known, it is thought that the polymer might bind to DNA and inhibit the action of nucleases and/or bind to cells and promote endocytosis of the DNA.

Transfection mediated by DEAE-dextran differs from calcium phosphate coprecipitation in three important respects. First, it is generally used only for transient expression of cloned genes and not for stable transformation of cells. Second, it works very efficiently with lines of cells such as BSC-1, CV-1, and COS but is unsatisfactory with many other types of cells, perhaps because the polymer is toxic. Third, smaller amounts of DNA are used for transfection with DEAE-dextran than with calcium phosphate coprecipitation. Maximal transfection efficiency of 10^5 simian cells is achieved with 100–200 ng of supercoiled plasmid DNA; larger amounts of DNA (>2–3 μg) can be inhibitory. In contrast to transfection mediated by calcium phosphate, where high concentrations of DNA are required to promote the formation of a precipitate, carrier DNA is never used with the DEAE-dextran transfection method.

Since the method was introduced over 20 years ago, many variants of DEAE-dextran transfection have been described. In most cases, the differences are minor and reflect more the preferences of individual laboratories than fundamental improvements to the technique. However, there are two important variables that greatly affect the efficiency of the method: the concentration of DEAE-dextran that is used and the length of time that the cells are exposed to the DNA/DEAE-dextran mixture. It is possible to use either a relatively high concentration of DEAE-dextran (1 mg/ml) for short periods (30 minutes to 1.5 hours) or a lower concentration (250 μg/ml) for longer periods of time (up to 8 hours). The first of these transfection procedures is the more efficient, but it involves monitoring the cells for early signs of distress when they are exposed to the facilitator. This requires practice. The second technique is less stringent and is therefore more reliable.

Transfection Using DEAE-Dextran: Protocol I

1. Prepare the following solutions:

50 mg/ml DEAE-dextran

Dissolve 100 mg of DEAE-dextran (M_r = 500,000; Pharmacia) in 2 ml of distilled H_2O. Sterilize the solution by autoclaving for 20 minutes at 15 lb/sq. in. on liquid cycle. (*Note:* Autoclaving also assists dissolution of the polymer.)

The molecular weight of the DEAE-dextran originally used for transfection was $> 2 \times 10^6$ (McCutchan and Pagano 1968). Although this material is no longer available commercially, it is still occasionally found in chemical storerooms. The older batches of higher-molecular-weight DEAE-dextran are more efficient facilitators of transfection than those currently available.

Tris-buffered saline–dextrose (TBS-D)

See Appendix B for preparation of TBS. Immediately before use, add 1 ml of a 20% (w/v) solution of dextrose in water (autoclaved) to each aliquot of TBS. The final concentration of dextrose should be 0.1%.

Phosphate-buffered saline (PBS)

See Appendix B for preparation of PBS.

2. Twenty-four hours before transfection, harvest exponentially growing cells by trypsinization and replate 60-mm tissue culture dishes with 10^5 cells/dish (or 35-mm dishes with 5×10^4 cells/dish). Add 5 ml (or 3 ml for 35-mm dish) of complete growth medium, and incubate the cultures for 20–24 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

The cells should be approximately 75% confluent at the time of transfection. If the cells are grown for less than 12 hours prior to transfection, they will be less well anchored to the substratum and more likely to detach during exposure to DEAE-dextran.

3. Prepare the DNA/DEAE-dextran/TBS-D solution (250 μl per 60-mm dish; 150 μl per 35-mm dish) by mixing 0.1–4 μg/ml supercoiled or circular DNA and 1.0 mg/ml DEAE-dextran in TBS-D.

The amount of DNA required to achieve maximal levels of transient expression depends on the exact nature of the construct and should be determined in preliminary experiments. If the construct carries a replicon that will function in the transfected cells, 100–200 ng of DNA per 10^5 cells should be sufficient; if no replicon is present, larger amounts of DNA may be required (up to 2 μg per 10^5 cells).

4. Remove the medium by aspiration, and wash the monolayers twice with prewarmed (37°C) PBS and once with prewarmed TBS-D.

5. Add the DNA/DEAE-dextran/TBS-D solution (250 μl per 60-mm dish;

150 μl per 35-mm dish). Rock gently to spread the solution evenly across the monolayer. Return the cultures to the incubator for 30–90 minutes (the time will depend on the sensitivity of each batch of cells to the DNA/DEAE-dextran/TBS-D solution). At 15–20-minute intervals, remove the dishes from the incubator, swirl them gently, and check the appearance of the cells under the microscope. If the cells are still firmly attached to the substratum, continue the incubation. The incubation should be halted when the cells begin to shrink and round up.

6. Remove the DNA/DEAE-dextran/TBS-D solution by aspiration. Gently wash the monolayers once with prewarmed (37°C) TBS-D and then once with prewarmed PBS, taking care not to dislodge the transfected cells.

7. Add 5 ml (per 60-mm dish) or 3 ml (per 35-mm dish) of prewarmed (37°C) medium supplemented with serum and chloroquine diphosphate (100 μM final concentration), and incubate the cultures for 3–5 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

   The efficiency of transfection is increased severalfold by treatment with chloroquine, which may act by inhibiting the degradation of the DNA by lysosomal hydrolases (Luthman and Magnusson 1983). Chloroquine diphosphate is stored as a sterile stock solution (100 mM; 60 mg/ml in water) in foil-wrapped tubes at −20°C. Note, however, that the cytotoxic effects of a combination of DEAE-dextran and chloroquine can be severe. It is therefore important to carry out preliminary experiments to determine the maximum permissible length of exposure to chloroquine after treatment of cells with DEAE-dextran.

8. Remove the medium by aspiration, and wash the monolayers three times with prewarmed (37°C) serum-free medium. Add to the cells 5 ml (per 60-mm dish) or 3 ml (per 35-mm dish) of medium supplemented with serum, and incubate the cultures for 36–60 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2 before assaying for expression of the transfected DNA. Depending on the experiment, continue with step 9a (transient expression) or 9b (stable transformation).

9. a. *Transient expression:* Harvest the cells 48–60 hours after transfection for analysis of RNA or DNA by hybridization. Newly synthesized protein may be analyzed by radioimmunoassay, by western blotting, by immunoprecipitation following in vivo metabolic labeling, or by assays of enzymatic activity in cell extracts. For assays that involve replicate samples or treatment of transfected cells under multiple conditions or over a time course, it is desirable to avoid dish-to-dish variation in transfection efficiency. In these cases, it is best to transfect large monolayers of cells (90-mm dishes) and then to trypsinize the cells after 24 hours of incubation and distribute them among several smaller dishes.

   b. *Stable transformation:* Following 18–24 hours of incubation in nonselective medium to allow expression of the transferred gene(s) to occur, the cells are trypsinized and replated in the appropriate selective medium. This medium should be changed every 2–4 days for 2–3

weeks to remove the debris of dead cells and to allow colonies of resistant cells to grow.

Individual colonies may be cloned and propagated for assay (for methods, see Jakoby and Pastan 1979). A permanent record of the numbers of colonies may be obtained by fixing the remaining cells with ice-cold methanol for 15 minutes and then staining them with 10% Giemsa for 15 minutes at room temperature before rinsing in tap water. The Giemsa stain should be freshly prepared in PBS or water and filtered through Whatman No. 1 filter paper.

The dilution at which the cells should be replated to yield well-separated colonies will be determined by the efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection.

Transfection Using DEAE-Dextran: Protocol II

1. Prepare the following solutions:

50 mg/ml DEAE-dextran

Dissolve 100 mg of DEAE-dextran (M_r = 500,000; Pharmacia) in 2 ml of distilled H_2O. Sterilize the solution by autoclaving for 20 minutes at 15 lb/sq. in. on liquid cycle. (*Note:* Autoclaving also assists dissolution of the polymer.)

The molecular weight of the DEAE-dextran originally used for transfection was $>2\times10^6$ (McCutchan and Pagano 1965). Although this material is no longer available commercially, it is still occasionally found in chemical storerooms. The older batches of higher-molecular-weight DEAE-dextran are more efficient facilitators of transfection than those currently available.

HEPES-buffered Dulbecco's modified Eagle's medium (HEPES-buffered DMEM)

DMEM lacking $NaHCO_3$ but containing 10 mM HEPES (pH 7.15)

No serum should be added to this reagent.

2. Twenty-four hours before transfection, harvest exponentially growing cells by trypsinization and replate 60-mm tissue culture dishes with 10^5 cells/dish (or 35-mm dishes with 5×10^4 cells/dish). Add 5 ml (or 3 ml for 35-mm dish) of complete growth medium, and incubate the cultures for 20–24 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

The cells should be approximately 75% confluent at the time of transfection. If the cells are grown for less than 12 hours prior to transfection, they will be less well anchored to the substratum and more likely to detach during exposure to DEAE-dextran.

3. Prepare the DNA/DEAE-dextran solution (500 μl per 60-mm dish; 250 μl per 35-mm dish) by mixing 0.1–4 μg/ml supercoiled or circular DNA and 250 μg/ml DEAE-dextran in HEPES-buffered DMEM.

The amount of DNA required to achieve maximal levels of transient expression depends on the exact nature of the construct and should be determined in preliminary experiments. If the construct carries a replicon that will function in the transfected cells, 100–200 ng of DNA per 10^5 cells should be sufficient; if no replicon is present, larger amounts of DNA may be required (up to 2 μg per 10^5 cells).

4. Remove the medium from the cells by aspiration, and wash the monolayers twice with prewarmed (37°C) HEPES-buffered DMEM. This step is important for maximal efficiency of the procedure.

5. Add the DNA/DEAE-dextran/DMEM solution to the cells (0.5 ml per 60-mm dish; 0.25 ml per 35-mm dish), and return the cells to the incubator for up to 8 hours. Gently rock the dishes every 2 hours to ensure even exposure to the DNA/DEAE-dextran/DMEM solution.

The efficiency of transfection is increased severalfold by concurrent treatment of the cells with chloroquine diphosphate, which may act by inhibiting the degradation of the DNA by lysosomal hydrolases (Luthman and Magnusson 1983). The drug (final concentration 100 μM) is added to the DNA/DEAE-dextran solution just before it is applied to the cells. Because chloroquine is toxic, the time of incubation must then be limited to 3–5 hours. Chloroquine diphosphate is stored as a sterile stock solution (100 mM; 60 mg/ml in water) in foil-wrapped tubes at −20°C.

6. Remove the DNA/DEAE-dextran/DMEM solution from the cells by aspiration, and gently wash the cell monolayers twice with prewarmed (37°C) HEPES-buffered DMEM. Take care not to dislodge the transfected cells. Wash the cells once with prewarmed DMEM (buffered with $NaHCO_3$ and not HEPES) supplemented with serum. Add to the cells 5 ml (per 60-mm dish) or 3 ml (per 35-mm dish) of complete growth medium, and incubate the cultures for 36–60 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2 before assaying for expression of the transfected DNA. Depending on the experiment, continue with step 7a (transient expression) or step 7b (stable transformation).

7. a. *Transient expression:* Harvest the cells 48–60 hours after transfection for analysis of RNA or DNA by hybridization. Newly synthesized protein may be analyzed by radioimmunoassay, by western blotting, by immunoprecipitation following in vivo metabolic labeling, or by assays of enzymatic activity in cell extracts. For assays that involve replicate samples or treatment of transfected cells under multiple conditions or over a time course, it is desirable to avoid dish-to-dish variation in transfection efficiency. In these cases, it is best to transfect large monolayers of cells (90-mm dishes) and then to trypsinize the cells after 24 hours of incubation and distribute them among several smaller dishes.

   b. *Stable transformation:* Following 18–24 hours of incubation in nonselective medium to allow expression of the transferred gene(s) to occur, the cells are trypsinized and replated in the appropriate selective medium. This medium should be changed every 2–4 days for 2–3 weeks to remove the debris of dead cells and to allow colonies of resistant cells to grow.

   Individual colonies may be cloned and propagated for assay (for methods, see Jakoby and Pastan 1979). A permanent record of the numbers of colonies may be obtained by fixing the remaining cells with ice-cold methanol for 15 minutes and then staining them with 10% Giemsa for 15 minutes at room temperature before rinsing in tap water. The Giemsa stain should be freshly prepared in PBS or water and filtered through Whatman No. 1 filter paper.

The dilution at which the cells should be replated to yield well-separated colonies will be determined by the efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection.

DNA TRANSFECTION USING POLYBRENE

The polycation Polybrene has been used as a facilitator of DNA transfection into cells that have proved to be relatively resistant to transfection using calcium phosphate coprecipitation (Kawai and Nishizawa 1984; Chaney et al. 1986). The method outlined below works efficiently for stable transformation of CHO cells by plasmid DNA, yielding approximately 15-fold more transformants than calcium phosphate–DNA coprecipitation. However, there is no difference between the two methods in the efficiency of transformation of high-molecular-weight DNA. It is not known whether Polybrene-mediated transfection can be used for transient expression of cloned DNA or whether it can be adapted for stable transformation with cell lines other than CHO.

1. Harvest exponentially growing CHO cells by trypsinization, and replate them at a density of 5×10^5 cells per 90-mm tissue culture dish in 10 ml of alpha medium (GIBCO) containing 10% fetal calf serum. Incubate the cultures for 18–20 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

2. Replace the medium with 3 ml of serum-containing medium containing DNA (5 ng to 40 μg; no carrier DNA) and 30 μg of Polybrene (Aldrich). Mix the DNA with the medium before adding the Polybrene solution (10 mg/ml in water and sterilized by passage through a 0.22-micron filter). Return the cells to the incubator for 6 hours, gently rocking the dishes every 90 minutes to ensure even exposure of the cells to the DNA–Polybrene mixture.

3. Remove the medium containing the DNA and Polybrene by aspiration, and treat the cells with dimethyl sulfoxide (DMSO) for 4 minutes by adding 5 ml of a 30% solution of DMSO in serum-containing alpha medium. Remove the DMSO solution by aspiration, wash the cells once with prewarmed (37°C), serum-free medium, and add 10 ml of complete medium containing 10% fetal calf serum. Incubate the cultures for 48 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

4. Recover the cells by trypsinization, and replate them in appropriate selective medium. Change the medium every 2–4 days for 2–3 weeks to allow death of sensitive cells and the emergence of resistant colonies.

   The dilution at which the cells should be replated to yield well-separated colonies is determined by the efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection. Typically, 1 μg of a standard plasmid such as pSV2-*neo* will transform approximately 3% of the cells to neomycin resistance.

5. Individual colonies may be cloned and propagated for assay (for methods, see Jakoby and Pastan 1979). A permanent record of the numbers of colonies may be obtained by fixing the remaining cells as described in step 7b, page 16.46.

DNA TRANSFECTION BY PROTOPLAST FUSION

Cloned DNA can be introduced into mammalian cells by fusing protoplasts, prepared from bacteria carrying the plasmid DNA of interest, with cultured cells (Schaffner 1980; Rassoulzadegan et al. 1982). The bacteria are grown in the presence of chloramphenicol to amplify the plasmid DNA and then treated with lysozyme to remove the cell wall. The resulting protoplasts are centrifuged onto a monolayer of mammalian cells, and the resulting mixture is treated with polyethylene glycol (PEG) to promote fusion. During this process, bacterial and plasmid DNAs are transferred into the mammalian cell. PEG is then removed, and the cells are incubated in fresh tissue culture medium containing kanamycin to inhibit the growth of any surviving bacteria.

Protoplast fusion has been used both for transient expression of cloned genes and for establishment of stable lines of mammalian cells. An estimate of the efficiency of these processes can be gained from the following data. When bacteria and mammalian cells were mixed in a ratio of 10,000:1, approximately 6% of the mammalian cells transiently expressed a gene carried on the amplified plasmid (Schaffner 1980). In subsequent experiments, in which a different bacterial host cell was employed, 100% of the mammalian cells were transiently transfected (Rassoulzadegan et al. 1982). The efficiency of stable transformation achieved in this experiment was 0.02% (40 transformants per 200,000 mammalian cells, with an input ratio of 10,000 protoplasts per mammalian cell).

Protoplast fusion has been used to stably introduce immunoglobulin genes into B cells (Gillies et al. 1983) and globin genes into mouse erythroleukemia cells (Charnay et al. 1984). The advantage of this method is its high efficiency. However, the manipulations are time-consuming and cotransformation is usually not possible. Thus, the gene of interest must always be carried on a plasmid containing the desired selectable marker.

Preparation of Protoplasts

1. Inoculate 100 ml of LB medium containing the appropriate antibiotic with a fresh overnight culture of bacteria containing the plasmid of interest. When the bacteria have grown to mid-log phase ($\sim 2 \times 10^8$ bacteria/ml), chloramphenicol should be added from a stock solution (34 mg/ml in absolute ethanol and sterilized by passage through a 0.22-micron filter) to a final concentration of 250 μg/ml. Continue incubation overnight.

   See Chapter 1, pages 1.21 and 1.33, for a discussion of the use of chloramphenicol in the amplification of plasmids.

2. Recover the bacterial cells from 50 ml of the culture by centrifugation at 3000*g* for 10 minutes at 4°C, and resuspend the pellet in 2.5 ml of an ice-cold solution of 20% (w/v) sucrose in 50 mM Tris · Cl (pH 8.0).

3. Add 0.5 ml of a fresh solution of lysozyme (5 mg/ml in 250 mM Tris · Cl [pH 8.0]). Incubate the suspension for 5 minutes at 4°C.

   Lysozyme will not work efficiently if the pH of the solution is less than 8.0.

4. Add 1 ml of ice-cold 0.25 M EDTA (pH 8.0). Incubate on ice for 5 minutes.

5. Slowly add 1 ml of ice-cold 50 mM Tris · Cl (pH 8.0), and incubate the suspension for 15 minutes at 37°C. The bacterial cell wall is digested during this incubation. The resulting protoplasts are fragile and must be treated gently in all subsequent steps.

   The efficiency of protoplast formation varies from one strain of bacteria to another and is affected by the physiological state of the cells. It is therefore essential to monitor the appearance of protoplasts under a microscope. Prior to protoplast formation, the bacteria should appear as rod-shaped particles. When the peptidoglycan layer surrounding the inner membrane has been digested by lysozyme, the cells become spherical. Digestion should be stopped when approximately 80% of the bacteria have been converted to protoplasts.

6. Slowly add 20 ml of 10% (w/v) sucrose, 10 mM $MgCl_2$ made up in Dulbecco's modified Eagle's medium (DMEM) and equilibrated at 37°C. Gently swirl the tube to mix the solution. Incubate the suspension for 15 minutes at room temperature. Check the protoplasts using the following criteria:

   a. The color should be red-orange.

   b. Upon swirling, there should be turbidity but the viscosity should be low.

   c. Debris, if present at all, should be minimal.

   The final concentration of protoplasts is approximately 5×10^8/ml.

   This step is included to eliminate the viscosity that results from the liberation of high-molecular-weight DNA from the small percentage of protoplasts that lyse during preparation.

Fusion of Protoplasts to Adherent Mammalian Cells

1. Twenty-four hours before transfection, harvest exponentially growing cells by trypsinization and seed six-well tissue culture plates with 7.5×10^5 cells/well. Incubate the cultures overnight at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2.

2. Remove the medium by aspiration, rinse the wells once with DMEM, and add 1.5 ml of the suspension of bacterial protoplasts prepared in step 6, page 16.49. This corresponds to approximately 10,000 protoplasts for every mammalian cell. Deposit the protoplasts onto the cells by centrifugation at 500*g* for 10 minutes at room temperature in a IEC-8R centrifuge (rotor no. 216) (or its equivalent).

   The protoplast fusion can also be carried out in 60-mm dishes, but plates are more convenient to centrifuge. Use approximately 10,000 protoplasts for every mammalian cell. When centrifuging bacterial protoplasts onto the adherent layer of mammalian cells in 60-mm dishes, it is best to stack the dishes in an empty tissue culture dish and tape the stack together. The stack is then placed on the flat bottom of a swinging-bucket rotor for centrifugation. A few sheets of tissue (Kimwipes or Kleenex) should be placed at the bottom of the stack to prevent the bottom dish from cracking.

3. Immediately after centrifugation, remove the supernatant from the wells by aspiration without tilting the plate. Add 2 ml of 50% polyethylene glycol 1000 (PEG 1000) in DMEM minus fetal calf serum (equilibrated to 37°C) above the same spot used for aspiration. Incubate the plates for 1–2 minutes at room temperature.

   In a microwave oven, melt sufficient PEG 1000 to make a 50% solution by volume in DMEM (prewarmed to 37°C). Sterilize the solution by passage through a 0.22-micron filter.

   Different cell lines may require different times of incubation or different concentrations of PEG to achieve the optimal balance between toxicity and efficient fusion. For example, 42.5% PEG is used to fuse bacterial protoplasts to WI-38 cells because higher concentrations of PEG are toxic. When using protoplast fusion for the first time with a particular line of mammalian cells, preliminary experiments should be carried out to measure the amount of transient expression obtained when fusion performed under sets of different conditions.

4. Add 5 ml of prewarmed (37°C) medium to each well at the same spot used for aspiration. Remove the medium from the plate by aspiration at the same spot. Do not agitate. Wash the monolayer of cells four times with prewarmed (37°C) DMEM without fetal calf serum.

   The scummy layer of protoplasts that adheres to the monolayer of cells should not be rinsed away.

5. Add the appropriate amount of prewarmed (37°C) tissue culture medium containing serum, penicillin (100 units/ml), streptomycin (100 μg/ml), and kanamycin (100 μg/ml). Incubate the dishes for 36–60 hours at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2. Depending on the experiment continue with step 6a (transient expression) or 6b (stable transformation).

6. a. *Transient expression:* Harvest the cells 48–60 hours after transfection for analysis of RNA or DNA by hybridization. Newly synthesiz d protein may be analyzed by radioimmunoassay, by western blotting, by immunoprecipitation following in vivo metabolic labeling, or by assays of enzymatic activity in cell extracts. For assays that involve replicate samples or treatment of transfected cells under multiple conditions or over a time course, it is desirable to avoid dish-to-dish variation in transfection efficiency. In these cases, it is best to transfect large monolayers of cells (90-mm dishes) and then to trypsinize the cells after 24 hours of incubation and distribute them among several smaller dishes.

b. *Stable transformation:* Following 18–24 hours of incubation in non-selective medium to allow expression of the transferred gene(s) to occur, the cells are trypsinized and replated in the appropriate selective medium. This medium should be changed every 2–4 days for 2–3 weeks to remove the debris of dead cells and to allow colonies of resistant cells to grow.

Individual colonies may be cloned and propagated for assay (for methods, see Jakoby and Pastan 1979). A permanent record of the numbers of colonies may be obtained by fixing the remaining cells with ice-cold methanol for 15 minutes and then staining them with 10% Giemsa for 15 minutes at room temperature before rinsing in tap water. The Giemsa stain should be freshly prepared in PBS or water and filtered through Whatman No. 1 filter paper.

The dilution at which the cells should be replated to yield well-separated colonies will be determined by the efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection.

Fusion of Protoplasts to Mammalian Cells Growing in Suspension

1. Recover the mammalian cells growing in suspension by centrifugation at 1000*g* for 5 minutes at 4°C. Wash the cells twice in ice-cold phosphate-buffered saline (PBS; see Appendix B).

2. Add to the mammalian cell pellet the protoplasts prepared in step 6, page 16.49. Use a ratio of protoplasts to mammalian cells of 10,000:1. Gently resuspend the mammalian cells in the suspension of protoplasts by drawing the mixture up and down in a wide-bore plastic pipette.

   Important: Do not vortex! If the protoplasts lyse at this stage and the solution becomes viscous, the efficiency of transformation will almost certainly be very low.

3. Centrifuge the mixture of bacterial protoplasts and mammalian cells at 2000 rpm for 10 minutes at room temperature in a Beckman TJ6 centrifuge (or its equivalent). Remove the supernatant by aspiration.

4. Flick the side of the tube with a finger until the pellet is loosened, and then add 0.5 ml of 50% polyethylene glycol (PEG 1000) in Dulbecco's modified Eagle's medium (DMEM) (equilibrated to 37°C). Incubate for 1 minute at room temperature.

   In a microwave oven, melt sufficient PEG 1000 to make a 50% solution by volume in DMEM (prewarmed to 37°C). Sterilize the solution by passage through a 0.22-micron filter.

5. Over a period of approximately 5 minutes while still at room temperature, dilute the suspension to 20 ml with DMEM or PBS.

6. Recover the cells by centrifugation at 1000*g* for 5 minutes at room temperature. Gently resuspend the pellet in prewarmed (37°C) medium containing serum, penicillin (100 units/ml), streptomycin (100 μg/ml), and kanamycin (100 μg/ml). Transfer the cells to a tissue culture dish and incubate at 37°C in a humidified incubator in an atmosphere of 5–7% CO_2. Depending on the experiment continue with step 7a (transient expression) or 7b (stable transformation).

7. a. *Transient expression:* Harvest the cells 48–60 hours after transfection for analysis of RNA or DNA by hybridization. Newly synthesized protein may be analyzed by radioimmunoassay, by western blotting, by immunoprecipitation following in vivo metabolic labeling, or by assays of enzymatic activity in cell extracts. For assays that involve replicate samples or treatment of transfected cells under multiple conditions or over a time course, it is desirable to avoid dish-to-dish variation in transfection efficiency. In these cases, it is best to transfect large numbers of cells (90-mm dishes) and then to divide the cells after 24 hours of incubation and distribute them among several smaller dishes.

   b. *Stable transformation:* Following 18–24 hours of incubation in non-selective medium to allow expression of the transferred gene(s) to occur, the cells are divided and plated in the appropriate selective medium.

This medium should be changed every 2–4 days for 2–3 weeks to remove the debris of dead cells and to allow colonies of resistant cells to grow.

The dilution at which the cells should be plated will be determined by th efficiency of stable transformation, which can vary over several orders of magnitude (see, e.g., Spandidos and Wilkie 1984). The efficiency is dependent on (1) the recipient cell type (significant differences have been observed even between different clones or different passage numbers of the same cell line [Corsaro and Pearson 1981; Van Pel et al. 1985]), (2) the nature of the introduced gene and the efficacy of the transcriptional control signals associated with it, and (3) the amount of donor DNA used in the transfection.

DNA TRANSFECTION BY ELECTROPORATION

The use of pulsed electric fields to introduce DNA into cells in culture has been termed electroporation (Neumann et al. 1982). The method has been used to introduce DNA into a variety of animal cells (Neumann et al. 1982; Wong and Neumann 1982; Potter et al. 1984; Sugden et al. 1985; Toneguzzo et al. 1986; Tur-Kaspa et al. 1986), plant cells (Fromm et al. 1985, 1986; Ecker and Davis 1986) and recently into bacteria. (For a review of methods for the introduction of DNA molecules into eukaryotic cells by electroporation, see Andreason and Evans [1988]. For a discussion of the use of electroporation to introduce DNA into bacterial cells, see Chapter 1, page 1.74.) The procedure has been used for both transient expression and stable transformation, but the efficiency of transfection has varied widely in reports from different laboratories. Recently, two groups have examined the parameters that affect transfection efficiencies and have shown that different conditions must be used for different cell lines (Chu et al. 1987; Knutson and Yee 1987). It is therefore essential to carry out a series of preliminary experiments to determine the conditions that lead to acceptable levels of transient expression or stable transformation for a particular cell line.

The efficiency of transfection by electroporation is influenced by a number of factors.

- *The strength of the applied electric field.* At low voltage, the plasma membranes of cultured cells are not sufficiently altered to allow passage of DNA molecules; at higher voltage, the cells are irreversibly damaged. For most lines of mammalian cells, the maximal level of transient expression (as measured by assays of CAT activity, for example) is reached when voltages between 250 V/cm and 750 V/cm are applied. Typically, between 20% and 50% of the cells survive this treatment (as measured by exclusion of trypan blue [Patterson 1979]).

- *The length of the electric pulse.* Usually, a single electric pulse is passed through the cells. Some electroporation devices allow the experimenter to control the length and shape of the pulse; in others, the characteristics of the pulse are determined solely by the capacitance of the power supply. The available data, although scanty, indicate that the optimal length of the electric pulse required for electroporation is 20–100 milliseconds. The efficiency of transient expression is increased if the cells are incubated for 1–2 minutes in the electroporation chamber after exposure to the electric pulse (Rabussay et al. 1987).

- *Temperature.* Some workers report that maximal levels of transient expression are obtained when the cells are maintained at room temperature during electroporation (Chu et al. 1987); others have obtained better results when the cells are maintained at 0°C (Reiss et al. 1986). These discrepancies may result from differences in the responses of various types of mammalian cells to the passage of electric current or in the amount of heat generated during electroporation when large electrical voltages (>1000 V/cm) and/or extended electric pulses (>100 milliseconds) are used.

- *Conformation and concentration of DNA.* Although both linear and circular DNAs can be transfected by electroporation, higher levels of both transient expression and stable transformation are obtained when linear DNA is used (Neumann et al. 1982; Potter et al. 1984; Toneguzzo et al. 1986). Effective transfection has been obtained with concentrations of DNA ranging from 1 μg/ml to 40 μg/ml.
- *Ionic composition of the medium.* The efficiency of transfection is manyfold higher when the cells are suspended in buffered salt solutions (e.g., HEPES-buffered saline) rather than in buffered solutions of nonionic substances such as mannitol or sucrose (Rabussay et al. 1987).

Electroporation has one major advantage: It works well with cell lines that are refractive to other techniques, such as calcium phosphate–DNA coprecipitation. However, considerable work may be required to define optimal conditions for the particular cell line under study. Furthermore, the conditions established in one laboratory do not necessarily work well in another. A number of different electroporation instruments are commercially available, and the manufacturers provide detailed protocols for their use.

Strategies for Studying Gene Regulation

Our understanding of the ways in which mammalian genes are regulated comes chiefly from experiments in which the activity of wild-type and mutant versions of putative *cis*-acting controlling elements is measured in transfected mammalian cells.

The accuracy with which transcripts are initiated can be determined by mapping the precise location of the 5′ terminus of the mRNA on its template DNA. Techniques to map the 5′ termini of mRNAs, which are described in Chapter 7, include primer extension, digestion of DNA:RNA hybrids with nuclease S1, and digestion of RNA:RNA hybrids with RNAase A. The last two methods have several advantages. These techniques can easily be adapted to yield estimates of the concentration of the RNA of interest. Furthermore, because only accurately initiated RNA is detected, any expression from cryptic promoters located in the flanking DNA sequences or the vector is ignored. By analyzing the RNA, it is also possible to avoid the use of a "reporter" gene. Thus, *cis*-acting regulatory sequences located within the gene of interest or downstream from it can be detected. However, it is often necessary to discriminate between transcripts originating from the transfected gene and those originating from the homologous gene present in the cellular chromosome. This is usually accomplished by deleting or inserting a small sequence in the 5′ untranslated region of the transfected DNA, so that the resulting transcript is readily distinguished from its normal counterpart. The transcript of the endogenous gene can then be used as an internal control in transfection experiments.

Direct assays of the abundance and structure of mRNAs have one major disadvantage: They quickly become tiresome, since RNAs encoded by many different mutants must be purified and analyzed by hybridization and gel electrophoresis. In most cases, therefore, the ability of the *cis*-acting elements to modulate the rate of transcription is not assayed directly. Instead, the *cis*-acting elements are joined to a reporter gene that codes for a novel enzymatic activity (e.g., chloramphenicol acetyltransferase [CAT]). The amount of the enzymatic activity that accumulates during the course of a transfection is taken as a measure of the ability of the *cis*-acting element to regulate transcription. Although this is an extremely indirect assay of the activity of *cis*-acting elements that control transcription, it is often the only one that is practicable. The major problem is that CAT activity can be generated not only from transcripts accurately initiated at the promoter of interest, but also from transcripts initiated at other sites in the vector (Langner et al. 1986). The level of such aberrant transcription may greatly dilute the contribution of the promoter being tested. Recently, Heard et al. (1987) demonstrated that the background caused by spurious transcription could be eliminated by placing a putative transcription termination signal upstream of the promoter. Such a signal can be provided by a DNA fragment located upstream of the murine c-*mos* gene (designated UMS for upstream mouse sequence). The mechanism by which UMS blocks read-through transcription is not known. Conceivably, the element contains a transcription terminator or destabilizes read-through mRNAs. To reduce the possibility of error caused by read-through mRNAs, the conclusions obtained by

analysis of mutant promoters/enhancers with reporter genes should be confirmed by direct assay of the mRNAs encoded by critical mutants.

VECTORS CARRYING REPORTER GENES

A number of different prokaryotic genes have been used as reporters of the transcriptional activity of mammalian promoters. However, not all of these are suitable for use in transient expression systems. For example, the bacterial *gpt* (Mulligan and Berg 1980), *neo*^{r} (Schöler and Gruss 1984), and *gal*K (galactokinase) (Schümperli et al. 1982) genes can be expressed in mammalian cells when linked to a mammalian promoter. However, the assays for all three of these enzymes are cumbersome, requiring, for example, the use of starch gel electrophoresis to separate the bacterial enzymatic activity from activities endogenous to mammalian cells.

The following considerations are important when selecting a reporter gene (Gorman et al. 1982a):

- The enzymatic activity encoded by the prokaryotic gene must be readily distinguishable from any similar activities present in the mammalian cells prior to transfection
- There should be no interference or competition from other enzymatic activities in the cells
- The assay for the encoded enzymatic activity should be rapid, sensitive, reproducible, and convenient

The enzyme CAT fulfills all of these criteria and its gene has therefore become the most widely used reporter gene for indirect assay of promoter activity in transfected mammalian cells. The *cat* gene was originally derived from the transposable element Tn*9* and confers resistance to chloramphenicol. The coding region is 1102 bp in length, which is ordinarily flanked by two 768-bp IS elements. A plasmid, pSV2CAT, has been constructed that contains the SV40 promoter/enhancer, 29 bp of 5′ untranslated sequence, the CAT coding sequence, and 8 bp of DNA 3′ to the UAA stop codon. pSV2CAT cannot confer chloramphenicol resistance on bacteria because the *cat* gene is not linked to a prokaryotic promoter. To assay putative promoters in mammalian cells, a derivative of pSV2CAT has been constructed (pSV0CAT; Gorman et al. 1982a) in which the promoter region of SV40 is replaced by the promoter being tested. CAT modifies and inactivates chloramphenicol by mono- and diacetylation, and a number of convenient assays have been developed to measure CAT activity in mammalian cells transfected with derivatives of pSV0CAT, including:

- *Incubation of extracts prepared from transfected cells with ^{14}C-labeled chloramphenicol.* The extent of modification of chloramphenicol is measured by thin-layer chromatography on silica gels (see pages 16.60–16.62), which separates the mono- and diacetylated derivatives of chloramphenicol from the unmodified compound. The silica gel is exposed to X-ray film and aligned with the resulting autoradiograph. Regions of the gel corresponding to spots on the film are scraped from the plate, and the amount of radioactivity is measured in a liquid scintillation counter. This assay can be somewhat tedious when large numbers of samples are assayed.

- *Incubation of extracts prepared from transfected cells with unlabeled chloramphenicol and ^{14}C-labeled acetyl coenzyme A.* CAT catalyzes the transfer of the ^{14}C-labeled acetyl group from acetyl coenzyme A to chloramphenicol. At the end of the reaction, the mixture is extracted with ethyl acetate. Acetylated forms of chloramphenicol partition into the organic phase, whereas the acetyl coenzyme A remains in the aqueous phase. The amount of chloramphenicol converted to the acetylated form can then be measured in a liquid scintillation counter. Before using this assay to measure CAT activity, it is necessary to heat the cell extracts for 10 minutes at 65°C to destroy an activity that consumes acetyl coenzyme A. This problem is not encountered with the silica gel assay because large amounts of unlabeled acetyl coenzyme A are present in the reaction.

When measuring the effect of promoters and enhancers on gene expression, it is essential to include an internal control that will distinguish differences in the level of transcription from differences in the efficiency of transfection or in the preparation of extracts. This is best achieved by cotransfecting the cells with two plasmids—one that carries the construct under investigation and another that constitutively expresses an activity that can be assayed in the same cell extracts prepared for measurement of CAT activity. An enzyme frequently used for this purpose is *E. coli* β-galactosidase, which is expressed in transfected mammalian cells from a promoter with a broad host range (the SV40 early promoter or the Rous sarcoma virus LTR). Extracts of most types of mammalian cells contain relatively low levels of endogenous β-galactosidase activity, and an increase in enzyme activity of up to 100-fold can usually be detected during the course of a transfection. However, β-galactosidase should not be used as an internal control in certain specialized cells (e.g., gut epithelial cells) that express high levels of this activity.

A number of different approaches can be used to normalize the CAT activity to the β-galactosidase activity. For example, the amount of protein in individual extracts prepared from a series of transfected cells is measured, and the CAT and β-galactosidase assays (see below) are then carried out using a standard amount of protein in each assay. The CAT activity is then normalized to the β-galactosidase activity. Alternatively, the β-galactosidase activity in a constant volume of extract is measured and the CAT assay is then carried out with amounts of extract that contain a defined amount of β-galactosidase activity. Finally, both enzymatic assays can be carried out in a constant volume of extract, and the results can then be normalized to a defined level of β-galactosidase activity.

The human growth hormone gene has also been used as an internal control for transfection (Selden et al. 1986). In this case, the coding region of the growth hormone gene is linked to a metallothionein or to the Rous sarcoma virus LTR—promoters that are expressed in most types of cultured mammalian cells. Rather than assaying extracts of transfected cells, the amount of growth hormone secreted into the tissue culture medium is analyzed using a commercially available radioimmunoassay. Although this assay is simple to carry out, it should be used with caution, since it is based on the assumption that the appearance of secreted protein in the medium parall ls the accumulation of a cytoplasmic enzyme. Thus, the assay does not necessarily provide an accurate internal control for the measurement of enzymatic activity in cell extracts.

ASSAYS FOR CHLORAMPHENICOL ACETYLTRANSFERASE AND β-GALACTOSIDASE ACTIVITIES

As discussed above, the best method to evaluate and optimize the transf c-tion efficiency of a particular cell line by a particular method is to use plasmids containing the gene for chloramphenicol acetyltransferase (CAT) or β-galactosidase. The time after transfection at which CAT or β-galactosidase activity is maximal must usually be determined empirically. In most cases, maximal activity is detected between 48 and 60 hours after transfection.

Preparation of Extracts

1. Remove the medium from cells growing in monolayers in 90-mm tissue culture dishes by gentle aspiration. Wash the monolayers three times with 5 ml of phosphate-buffered saline (PBS; see Appendix B) lacking magnesium and calcium ions.

2. Stand the dishes at an angle for 2–3 minutes to allow the last traces of PBS to drain to one side. Remove the last traces of PBS by aspiration. Add 1 ml of PBS to each plate, and using a policeman, scrape the cells into microfuge tubes. Store the microfuge tubes on ice until all of the plates have been processed.

3. Recover the cells by centifugation at 12,000*g* for 10 seconds at room temperature in a microfuge. Gently resuspend the cell pellets in 1 ml of ice-cold PBS, and again recover the cells by centrifugation. Remove the last traces of PBS from the cell pellets and from the walls of the microfuge tubes.

   The PBS can be conveniently removed with a disposable pipette tip attached to a vacuum line. Use gentle suction, and touch the tip to the surface of the liquid. Keep the tip as far away from the cell pellet as possible while the fluid is withdrawn from the tube. The pipette tip can then be used to vacuum the walls of the tube to remove any adherent droplets of fluid.

   At this stage, the cell pellet can be stored at −20°C for future analysis.

Assays for Chloramphenicol Acetyltransferase

METHOD 1: THIN-LAYER CHROMATOGRAPHY

This is a modification of methods described by Gorman et al. (1982a).

1. Resuspend the cell pellet from one 90-mm dish in 100 μl of 0.25 M Tris·Cl (pH 7.8). Vortex vigorously to break up clumps of cells.

2. Disrupt the cells by three cycles of freezing in dry ice and ethanol and thawing at 37°C. Make sure the tubes have been marked with ethanol-insoluble ink.

3. Centrifuge the suspension of disrupted cells at 12,000*g* for 5 minutes at 4°C in a microfuge. Transfer the supernatant to a fresh microfuge tube. Reserve 50 μl for the CAT assay, and store the remainder of the extract at −20°C.

4. Incubate the 50-μl aliquot of the extract for 10 minutes at 65°C to inactivate deacetylases. If the extract is cloudy or opaque at this stage, remove the particulate material by centrifugation at 12,000*g* for a further 2 minutes at 4°C in a microfuge.

   If the expression level is high, this incubation is not required (see page 16.58 and step 6 below).

5. Prepare CAT reaction mixture 1:

1 M Tris·Cl (pH 7.8)	50 μl
^{14}C-labeled chloramphenicol (60 mCi/mmole, diluted in water to 0.1 mCi/ml)	10 μl
acetyl coenzyme A (freshly prepared at a concentration of 3.5 mg/ml in H_2O)	20 μl

   For each 50 μl of cell extract to be tested, 80 μl of CAT reaction mixture 1 is required.

6. Mix each of the samples to be assayed with 80 μl of CAT reaction mixture 1, and incubate the reactions at 37°C. The length of the incubation depends on the concentration of CAT in the cell extract, which in turn depends on the strength of the promoter and the cell type under investigation. In most cases, incubation for 30 minutes to 2 hours is sufficient.

   The reactions can be incubated for longer periods of time (up to 16 hours) when the expression of the *cat* gene is low in the transfected cells. However, in this case, it is advisable to add another 10-μl aliquot of acetyl coenzyme A to each reaction after 2 hours of incubation. This is especially important if the extract has not been heated to 65°C to inactive enzymes that consume acetyl coenzyme A.

7. Add 1 ml of ethyl acetate to each sample, and mix thoroughly by vortexing for three periods of 10 seconds each. Centrifuge the mixtures at 12,000*g* for 5 minutes at room temperature in a microfuge.

The acetylated forms of chloramphenicol partition into the organic (upper) phase; unacetylated chloramphenicol remains in the aqueous phase.

8. Using an automatic pipettor, transfer exactly 900 μl of the upper phase to a fresh tube, carefully avoiding the lower phase and the interface. Discard the tube containing the lower phase in the radioactive waste.

9. Evaporate the ethyl acetate under vacuum. This is usually accomplished by placing the tubes in a rotating evaporator (e.g., Savant SpeedVac) for approximately 1 hour.

10. Redissolve the reaction products in 25 μl of ethyl acetate, carefully washing the sides of the tube.

11. Apply 10–15 μl of the dissolved reaction products to the origin of a 25-mm silica gel, thin-layer chromatography (TLC) plate. (The origin on the plate can be marked with a soft-lead pencil.) Apply 5 μl at a time, and evaporate the sample to dryness with a hair dryer after each application.

    A number of different TLC plates and solvents can be used to separate acetylated forms of chloramphenicol. We routinely use chloroform:methanol (95:5) and Sybron SIL G/UV254 (Brinkmann) TLC plates.

12. Prepare a TLC chamber containing 200 ml of chloroform:methanol (95:5).

13. Place the TLC plate in the chromatography chamber, close the chamber, and allow the solvent front to move approximately 75% of the distance to the top of the plate.

14. Remove the TLC plate from the tank and allow it to dry at room temperature. Place on the TLC plate adhesive dot labels marked with radioactive ink to align the plate with the film, and then expose the plate to X-ray film.

    Do not cover the TLC plate with Saran Wrap, which will block the relatively weak radiation emitted by ^{14}C.

    For preparation of radioactive ink, see Appendix E.

15. Develop the X-ray film and align it with the plate. Usually, three radioactive spots are visible. The spot that has migrated the least distance from the origin consists of nonacetylated chloramphenicol that partitioned into ethyl acetate. The two faster-migrating spots are modified forms of chloramphenicol that have been acetylated at one or the other of two potential sites. Diacetylated chloramphenicol may be detected as a third, even faster migrating spot only when high concentrations of CAT are used.

16. To quantitate CAT activity, cut the radioactive spots from the TLC plate and measure the amount of radioactivity they contain in a liquid scintillation counter (see Appendix E).

The sensitivity of the CAT assay can be optimized by eliminating sources of background. One source of background is due to read-through transcription—a problem that can be eliminated by placing a strong transcription termination element upstream of the promoter (Heard et al. 1987; Araki et al. 1988). "Intrinsic" acetylation of chloramphenicol in the absence of cellular extracts may also contribute to background. If this is a problem, decrease the concentration of chloramphenicol in the assay two- to fourfold (Heard et al. 1987).

METHOD 2: EXTRACTION WITH ORGANIC SOLVENTS

In this procedure (Sleigh 1986), radiolabeled acetyl coenzyme A is transferred to unlabeled chloramphenicol. The acetylated, radiolabeled forms of chloramphenicol are separated from unreacted acetyl coenzyme A by partitioning into ethyl acetate. The acetylated chloramphenicol partitions into the organic phase, and the acetyl coenzyme A remains in the aqueous phase.

Extracts of transfected cells, prepared as described on pages 16.59 and 16.60, steps 1–3 (except reserve 30 μl of the extract for the assay), are assayed as follows:

1. Prepare CAT reaction mixture 2:

8 mM chloramphenicol	20 μl
cell extract	30 μl
0.25 M Tris · Cl (pH 7.8)	30 μl
solution containing ^{14}C-labeled acetyl coenzyme A	20 μl

8 mM chloramphenicol (M_r = 321) is made by dissolving 2.57 mg of chloramphenicol in 1 ml of ethanol. The solution of ^{14}C-labeled acetyl coenzyme A is prepared by dissolving ^{14}C-labeled acetyl coenzyme A (58 Ci/mmole) in water to a concentration of 50 μCi/ml. Dispense the solution into 20-μl aliquots and store at −70°C. Just before setting up the reaction, dilute the ^{14}C-labeled acetyl coenzyme A tenfold into a solution of unlabeled acetyl coenzyme A (0.5 mM). The concentrations of the reagents in the final reaction mixture are: 2.3 mM chloramphenicol, 0.11 M Tris · Cl, and 129 μM (1 μCi/ml) ^{14}C-labeled acetyl coenzyme A.

For each 30 μl of cell extract to be tested, 70 μl of CAT reaction mixture 2 is required.

2. Add 30 μl of cell extract, prepared as described on page 16.59, to 70 μl of CAT reaction mixture 2. Incubate the reaction for 1 hour at 37°C and then for 10 minutes at 65°C to inactivate deacetylases.

3. Transfer the sample to an ice bath, and add 100 μl of ice-cold ethyl acetate. Mix the phases by vigorous vortexing. Centrifuge the mixtures at 12,000*g* for 3 minutes at room temperature in a microfuge.

4. Transfer 80 μl of the organic (upper) phase to a fresh tube. Add 100 μl of ice-cold ethyl acetate to the original tube (the aqueous phase). Vortex vigorously and centrifuge again as in step 3. Transfer 100 μl of the organic phase to the tube containing 80 μl of ethyl acetate from the first extraction.

Take great care not to transfer any of the aqueous phase, which contains radiolabeled substrate.

5. Mix the combined organic phases with 1 ml of scintillation fluid (Instagel, Hewlett-Packard, or equivalent) in a 1.5-ml plastic tube. Measure the radioactivity by liquid scintillation counting (see Appendix E).

Note

The linear range of the assay is between 0.03 and 0.25 unit of enzyme per milliliter of incubation mixture (a unit of CAT activity is defined as the amount of enzyme that transfers 1 nmole of acetyl groups from acetyl coenzyme A to chloramphenicol in 1 minute at 37°C [pH 7.8]).

gene be maintained intact and that the promoter region be active in the recipient cell. When total genomic DNA is used, efforts are made to keep the DNA as large as possible (usually >100 kb). For some genes (the 200-kb Factor VIII gene), even this large size may not be sufficient. One useful approach for many types of genes has been the use of restriction endonuclease digestion to determine sensitivity of the target gene to digestion with a particular enzyme. If the gene is not large, it is sometimes possible to prepare a genomic DNA library in a bacteriophage λ or cosmid vector and use purified library DNA for the transformation. Three advantages of this approach are:

1. The library can be divided into sublibraries and individual sublibraries retransformed. Once an active sublibrary is identified, that sublibrary is further subdivided and retransformed again. The process is repeated until a single clone is obtained.
2. It may be possible to enhance the expression of transfected genes by placing one or more enhancer-type elements into the vector. Such elements may improve the likelihood of a successful expression (see page 16.6).
3. In some cases, the vector sequences can be used to "rescue" a clone from the transfected cell line, if the vector sequences are maintained intact (Lindenmaier et al. 1982). Such an approach significantly speeds the process and eliminates the need to construct libraries from transfected cell lines.

The factors that determine whether a given promoter will be active in a given recipient cell line are not completely understood and are the subject of considerable research in their own right. To date, the source of the genomic DNA has been a factor in determining the transfectability of a given marker gene, except for trophoblast DNA, perhaps due to the high degree of methylation of such DNA. At some point, it may be possible to define what types of sequence elements (such as recipient-cell-specific enhancers) might improve the likelihood or level of expression of a foreign gene and thereby improve the success of cloning by genomic expression approaches.

References

Akusjärvi, G., C. Svensson, and O. Nygård. 1987. A mechanism by which adenovirus virus-associated RNA_I controls translation in a transient expression assay. *Mol. Cell. Biol.* **7:** 549.

Andreason, G.L. and G.A. Evans. 1988. Introduction and expression of DNA molecules in eukaryotic cells by electroporation. *BioTechniques* **6:** 650.

Araki, E., F. Shimada, M. Shichiri, M. Mori, and Y. Ebina. 1988. pSV00CAT: Low background CAT plasmid. *Nucleic Acids Res.* **16:** 1627.

Aruffo, A. and B. Seed. 1987. Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci.* **84:** 8573.

Bacchetti, S. and F.L. Graham. 1977. Transfer of the gene for thymidine kinase to thymidine kinase-deficient human cells by purified herpes simplex viral DNA. *Proc. Natl. Acad. Sci.* **74:** 1590.

Berkner, K.L. 1988. Development of adenovirus vectors for the expression of heterologous genes. *BioTechniques* **6:** 616.

Birnstiel, M.L., M. Busslinger, and K. Strub. 1985. Transcription termination and 3′ processing: The end is in site! *Cell* **41:** 349.

Boggs, S.S., R.G. Gregg, N. Borenstein, and O. Smithies. 1986. Efficient transformation and frequent single-site, single-copy insertion of DNA can be obtained in mouse erythroleukemia cells transformed by electroporation. *Exp. Hematol.* **14:** 988.

Bonthron, D.T., R.I. Handin, R.J. Kaufman, L.C. Wasley, E.C. Orr, L.M. Mitsock, B. Ewenstein, J. Loscalzo, D. Ginsburg, and S.H. Orkin. 1986. Structure of pre-pro-von Willebrand factor and its expression in heterologous cells. *Nature* **324:** 270.

Boshart, M., F. Weber, G. Jahn, K. Dorsch-Häsler, B. Fleckenstein, and W. Schaffner. 1985. A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* **41:** 521.

Broker, T.R. and M. Botchan. 1986. Papillomaviruses: Retrospectives and prospectives. *Cancer Cells* **4:** 17.

Campo, M.S. 1985. Bovine papillomavirus DNA: A eukaryotic cloning vector. In *DNA cloning: A practical approach* (ed. D.M. Glover), vol. 2, p. 213. IRL Press, Oxford.

Capecchi, M.R. 1980. High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell* **22:** 479.

Chaney, W.G., D.R. Howard, J.W. Pollard, S. Sallustio, and P. Stanley. 1986. High-frequency transfection of CHO cells using Polybrene. *Somatic Cell Mol. Genet.* **12:** 237.

Charnay, P., R. Treisman, P. Mellon, M. Chao, R. Axel, and T. Maniatis. 1984. Differences in human α- and β-globin gene expression in mouse erythroleukemia cells: The role of intragenic sequences. *Cell* **38:** 251.

Chen, C. and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7:** 2745.

———. 1988. Calcium phosphate-mediated gene transfer: A highly efficient transfection system for stably transforming cells with plasmid DNA. *BioTechniques* **6:** 632.

Chen, E.Y., P.M. Howley, A.D. Levinson, and P.H. Seeburg. 1982. The primary structure and genetic organization of the bovine papillomavirus type 1 genome. *Nature* **299:** 529.

Chia, W., M.R.D. Scott, and P.W.J. Rigby. 1982. The construction of cosmid libraries of eukaryotic DNA using the Homer series of vectors. *Nucleic Acids Res.* **10:** 2503.

Chu, G. and P.A. Sharp. 1981. SV40 DNA transfection of cells in suspension: Analysis of the efficiency of transcription and translation of T-antigen. *Gene* **13:** 197.

Chu, G., H. Hayakawa, and P. Berg. 1987. Electroporation for the efficient transfection of mammalian cells with DNA. *Nucleic Acids Res.* **15:** 1311.

Colbère-Garapin, F., F. Horodniceanu, P. Kourilsky, and A.-C. Garapin. 1981. A

sion of a transposable antibiotic resistance element in *Saccharomyces*. *Nature* **287:** 869.

Jolly, D.J., A.C. Esty, H.U. Bernard, and T. Friedmann. 1982. Isolation of a genomic clone partially encoding human hypoxanthine phosphoribosyltransferase. *Proc. Natl. Acad. Sci.* **79:** 5038.

Karin, M., G. Cathala, and M.C. Nguyen-Huu. 1983. Expression and regulation of a human metallothionein gene carried on an autonomously replicating shuttle vector. *Proc. Natl. Acad. Sci.* **80:** 4040.

Karlsson, S., K. Van Doren, S.G. Schweiger, A.W. Nienhuis, and Y. Gluzman. 1986. Stable gene transfer and tissue-specific expression of a human globin gene using adenoviral vectors. *EMBO J.* **5:** 2377.

Kaufman, R.J. 1987. High level production of proteins in mammalian cells. In *Genetic engineering: Principles and methods* (ed. J.K. Setlow), vol. 9, p. 155. Plenum Publishing, New York.

Kaufman, R.J. and P. Murtha. 1987. Translational control mediated by eucaryotic initiation factor-2 is restricted to specific mRNAs in transfected cells. *Mol. Cell. Biol.* **7:** 1568.

Kaufman, R.J. and P.A. Sharp. 1982a. Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary DNA gene. *J. Mol. Biol.* **159:** 601.

———. 1982b. Construction of a modular dihydrofolate reductase cDNA gene: Analysis of signals utilized for efficient expression. *Mol. Cell. Biol.* **2:** 1304.

Kaufman, R.J., P. Murtha, and M.V. Davies. 1987. Translational efficiency of polycistronic mRNAs and their utilization to express heterologous genes in mammalian cells. *EMBO J.* **6:** 187.

Kaufman, R.J., M.V. Davies, V.K. Pathak, and J.W.B. Hershey. 1989. The phosphorylation state of eucaryotic initiation factor 2 alters translational efficiency of specific mRNAs. *Mol. Cell. Biol.* **9:** 946.

Kaufman, R.J., P. Murtha, D.E. Ingolia, C.-Y. Yeung, and R.E. Kellems. 1986a. Selection and amplification of heterologous genes encoding adenosine deaminase in mammalian cells. *Proc. Natl. Acad. Sci.* **83:** 3136.

Kaufman, R.J., L.C. Wasley, B.C. Furie, B. Furie, and C.B. Shoemaker. 1986b. Expression, purification, and characterization of recombinant γ-carboxylated Factor IX synthesized in Chinese hamster ovary cells. *J. Biol. Chem.* **261:** 9622.

Kaufman, R.J., L.C. Wasley, A.J. Spiliotes, S.D. Gossels, S.A. Latt, G.R. Larsen, and R.M. Kay. 1985. Coamplification and coexpression of human tissue-type plasminogen activator and murine dihydrofolate reductase sequences in Chinese hamster ovary cells. *Mol. Cell. Biol.* **5:** 1750.

Kavathas, P. and L.A. Herzenberg. 1983. Stable transformation of mouse L cells for human membrane T-cell differentiation antigens, HLA and β_2-microglobulin: Selection by fluorescence-activated cell sorting. *Proc. Natl. Acad. Sci.* **80:** 524.

Kavathas, P., V.P. Sukhatme, L.A. Herzenberg, and J.R. Parnes. 1984. Isolation of the gene encoding the human T-lymphocyte differentiation antigen Leu-2 (T8) by gene transfer and cDNA subtraction. *Proc. Natl. Acad. Sci.* **81:** 7688.

Kawai, S. and M. Nishizawa. 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. *Mol. Cell. Biol.* **4:** 1172.

Kempe, T.D., E.A. Swyryd, M. Bruist, and G.R. Stark. 1976. Stable mutants of mammalian cells that overproduce the first three enzymes of pyrimidine nucleotide biosynthesis. *Cell* **9:** 541.

Kim, S.K. and B.J. Wold. 1985. Stable reduction of thymidine kinase activity in cells expressing high levels of antisense RNA. *Cell* **42:** 129.

Kioussis, D., F. Wilson, C. Daniels, C. Leveton, J. Taverne, and J.H.L. Playfair. 1987. Expression and rescuing of a cloned human tumour necrosis factor gene using an EBV-based shuttle cosmid vector. *EMBO J.* **6:** 355.

Kit, S., D.R. Dubbs, L.J. Piekarski, and T.C. Hsu. 1963. Deletion of thymidine kinase activity from L cells resistant to bromodeoxyuridine. *Exp. Cell Res.* **31:** 297.

Kitajewski, J., R.J. Schneider, B. Safer, S.M. Munemitsu, C.E. Samuel, B. Thimmappaya, and T. Shenk. 1986. Adenovirus VAI RNA antagonizes the antiviral action of interferon by preventing activation of the interferon-induced eIF-2α kinase. *Cell* **45:** 195.

Knutson, J.C. and D. Yee. 1987. Elec-

troporation: Parameters affecting transfer of DNA into mammalian cells. *Anal. Biochem.* **164:** 44.

Kozak, M. 1989. The scanning model for translation: An update. *J. Cell Biol.* **108:** 229.

Krainer, A.R. and T. Maniatis. 1988. RNA splicing. In *Transcription and splicing* (ed. B.D. Hames and D.M. Glover), p. 131. IRL Press, Oxford.

Krystal, M., R. Li, D. Lyles, G. Pavlakis, and P. Palese. 1986. Expression of the three influenza virus polymerase proteins in a single cell allows growth complementation of viral mutants. *Proc. Natl. Acad. Sci.* **83:** 2709.

Kühn, L.C., A. McClelland, and F.H. Ruddle. 1984. Gene transfer, expression, and molecular cloning of the human transferrin receptor gene. *Cell* **37:** 95.

Langner, K.-D., U. Weyer, and W. Doerfler. 1986. *Trans* effect of the E1 region of adenoviruses on the expression of a prokaryotic gene in mammalian cells: Resistance to 5′-CCGG-3′ methylation. *Proc. Natl. Acad. Sci.* **83:** 1598.

Law, M.-F., J.C. Byrne, and P.M. Howley. 1983. A stable bovine papillomavirus hybrid plasmid that expresses a dominant selective trait. *Mol. Cell. Biol.* **3:** 2110.

Lee, F., T. Yokota, T. Otsuka, L. Gemmell, N. Larson, J. Luh, K. Arai, and D. Rennick. 1985. Isolation of cDNA for a human granulocyte–macrophage colony-stimulating factor by functional expression in mammalian cells. *Proc. Natl. Acad. Sci.* **82:** 4360.

Lewis, J.A., K. Shimizu, and D. Zipser. 1983. Isolation and preliminary characterization of the Chinese hamster thymidine kinase gene. *Mol. Cell. Biol.* **3:** 1815.

Lin, P.-F., S.-Y. Zhao, and F.H. Ruddle. 1983. Genomic cloning and preliminary characterization of the human thymidine kinase gene. *Proc. Natl. Acad. Sci.* **80:** 6528.

Lindenmaier, W., H. Hauser, I. Greiser de Wilke, and G. Schütz. 1982. Gene shuttling: Moving of cloned DNA into and out of eukaryotic cells. *Nucleic Acids Res.* **10:** 1243.

Littlefield, J.W. 1964. Selection of hybrids from matings of fibroblasts in vitro and their presumed recombinants. *Science* **145:** 709.

———. 1966. The use of drug-resistant markers to study the hybridization of mouse fibroblasts. *Exp. Cell Res.* **41:** 190.

Liu, C.-C., C.C. Simonsen, and A.D. Levinson. 1984. Initiation of translation at internal AUG codons in mammalian cells. *Nature* **309:** 82.

Lowy, I., A. Pellicer, J.F. Jackson, G.-K. Sim, S. Silverstein, and R. Axel. 1980. Isolation of transforming DNA: Cloning the hamster *aprt* gene. *Cell* **22:** 817.

Luckow, V.A. and M.D. Summers. 1988. Trends in the development of baculovirus expression vectors. *Bio/Technology* **6:** 47.

Lupton, S. and A.J. Levine. 1985. Mapping genetic elements of Epstein-Barr virus that facilitate extrachromosomal persistence of Epstein-Barr virus-derived plasmids in human cells. *Mol. Cell. Biol.* **5:** 2533.

Lusky, M. and M. Botchan. 1981. Inhibition of SV40 replication in simian cells by specific pBR322 DNA sequences. *Nature* **293:** 79.

Luthman, H. and G. Magnusson. 1983. High efficiency polyoma DNA transfection of chloroquine treated cells. *Nucleic Acids Res.* **11:** 1295.

Mackett, M., G.L. Smith, and B. Moss. 1985. The construction and characterisation of vaccinia virus recombinants expressing foreign genes. In *DNA cloning: A practical approach* (ed. D.M. Glover), vol. 2, p. 191. IRL Press, Oxford.

Maniatis, T., S. Goodbourn, and J.A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science* **236:** 1237.

Mannino, R.J. and S. Gould-Fogerite. 1988. Liposome mediated gene transfer. *BioTechniques* **6:** 682.

Mansour, S.L., T. Grodzicker, and R. Tjian. 1985. An adenovirus vector system used to express polyoma virus tumor antigens. *Proc. Natl. Acad. Sci.* **82:** 1359.

Mason, P.J., J.A. Elkington, M.M. Lloyd, M.B. Jones, and J.G. Williams. 1986. Mutations downstream of the polyadenylation site of a *Xenopus* β-globin mRNA affect the position but not the efficiency of 3′ processing. *Cell* **46:** 263.

Maurer, R.A. 1989. Cationic liposome-mediated transfection of primary cultures of rat pituitary cells. *Bethesda Res. Lab. Focus* **11(2):** 25.

translated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46:** 659.

Shih, C. and R.A. Weinberg. 1982. Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* **29:** 161.

Shimizu, Y., B. Koller, D. Geraghty, H. Orr, S. Shaw, P. Kavathas, and R. DeMars. 1986. Transfer of cloned human class I major histocompatibility complex genes into HLA mutant human lymphoblastoid cells. *Mol. Cell. Biol.* **6:** 1074.

Simonsen, C.C. and A.D. Levinson. 1983. Isolation and expression of an altered mouse dihydrofolate reductase cDNA. *Proc. Natl. Acad. Sci.* **80:** 2495.

Simonsen, C.C., H.M. Shepard, P.W. Gray, D.W. Leung, D. Pennica, E. Yelverton, R. Derynck, P.J. Sherwood, A.D. Levinson, and D.V. Goeddel. 1982. Plasmid-directed synthesis of human immune interferon in *E. coli* and monkey cells. *UCLA Symp. Mol. Cell. Biol.* **25:** 1.

Sims, J.E., C.J. March, D. Cosman, M.B. Widmer, H.R. MacDonald, C.J. McMahan, C.E. Grubin, J.M. Wignall, J.L. Jackson, S.M. Call, D. Friend, A.R. Alpert, S. Gillis, D.L. Urdal, and S.K. Dower. 1988. cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin superfamily. *Science* **241:** 585.

Sleigh, M.J. 1986. A nonchromatographic assay for expression of the chloramphenicol acetyltransferase gene in eucaryotic cells. *Anal. Biochem.* **156:** 251.

Solnick, D. 1981. Construction of an adenovirus–SV40 recombinant producing SV40 T antigen from an adenovirus late promoter. *Cell* **24:** 135.

Southern, P.J. and P. Berg. 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Genet.* **1:** 327.

Spandidos, D.A. and L. Siminovitch. 1977. Transfer of codominant markers by isolated metaphase chromosomes in Chinese hamster ovary cells. *Proc. Natl. Acad. Sci.* **74:** 3480.

Spandidos, D.A. and N.M. Wilkie. 1984. Expression of exogenous DNA in mammalian cells. In *Transcription and translation: A practical approach* (ed. B.D. Hames and S.J. Higgins), p. 1. IRL Press, Oxford.

Stark, G.R. and G.M. Wahl. 1984. Gene amplification. *Annu. Rev. Biochem.* **53:** 447.

Subramani, S., R. Mulligan, and P. Berg. 1981. Expression of the mouse dihydrofolate reductase complementary deoxyribonucleic acid in simian virus 40 vectors. *Mol. Cell. Biol.* **1:** 854.

Sugden, B. 1982. Epstein-Barr virus: A human pathogen inducing lymphoproliferation *in vivo* and *in vitro*. *Rev. Infect. Dis.* **4:** 1048.

Sugden, B., K. Marsh, and J. Yates. 1985. A vector that replicates as a plasmid and can be efficiently selected in B-lymphoblasts transformed by Epstein-Barr virus. *Mol. Cell. Biol.* **5:** 410.

Summers, M.D. and G.E. Smith. 1987. A manual of methods for baculovirus vectors and insect cell culture procedures. *Tex. Agric. Exp. Stn. Bull.* No. 1555.

Szybalska, E.H. and W. Szybalski. 1962. Genetics of human cell lines, IV. DNA-mediated heritable transformation of a biochemical trait. *Proc. Natl. Acad. Sci.* **48:** 2026.

Thummel, C., R. Tjian, and T. Grodzicker. 1981. Expression of SV40 T antigen under control of adenovirus promoters. *Cell* **23:** 825.

———. 1982. Construction of adenovirus expression vectors by site-directed *in vivo* recombination. *J. Mol. Appl. Genet.* **1:** 435.

Thummel, C., R. Tjian, S.-L. Hu, and T. Grodzicker. 1983. Translational control of SV40 T antigen expressed from the adenovirus late promoter. *Cell* **33:** 455.

Toneguzzo, F., A.C. Hayday, and A. Keating. 1986. Electric field-mediated DNA transfer: Transient and stable gene expression in human and mouse lymphoid cells. *Mol. Cell. Biol.* **6:** 703.

Toole, J.J., J.L. Knopf, J.M. Wozney, L.A. Sultzman, J.L. Buecker, D.D. Pittman, R.J. Kaufman, E. Brown, C. Shoemaker, E.C. Orr, G.W. Amphlett, W.B. Foster, M.L. Coe, G.J. Knutson, D.N. Fass, and R.M. Hewick. 1984. Molecular cloning of a cDNA encoding human antihaemophilic factor. *Nature* **312:** 342.

Tooze, J., ed. 1980. *Molecular biology of tumor viruses,* 2nd edition: *DNA tumor viruses,* part 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Topp, W.C. 1981. Normal rat cell lines deficient in nuclear thymidine kinase. *Virology* **113:** 408.

Treisman, R., S.H. Orkin, and T. Maniatis. 1983. Specific transcription and RNA splicing defects in five cloned β-thalassaemia genes. *Nature* **302:** 591.

Treisman, R., U. Novak, J. Favaloro, and R. Kamen. 1981. Transformation of rat cells by an altered polyoma virus genome expressing only the middle-T protein. *Nature* **292:** 595.

Tur-Kaspa, R., L. Teicher, B.J. Levine, A.I. Skoultchi, and D.A. Shafritz. 1986. Use of electroporation to introduce biologically active foreign genes into primary rat hepatocytes. *Mol. Cell. Biol.* **6:** 716.

Urlaub, G. and L.A. Chasin. 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. *Proc. Natl. Acad. Sci.* **77:** 4216.

Vaheri, A. and J.S. Pagano. 1965. Infectious poliovirus RNA: A sensitive method of assay. *Virology* **27:** 434.

Van Doren, K., D. Hanahan, and Y. Gluzman. 1984. Infection of eucaryotic cells by helper-independent recombinant adenoviruses: Early region 1 is not obligatory for integration of viral DNA. *J. Virol.* **50:** 606.

Van Pel, A., E. De Plaen, and T. Boon. 1985. Selection of highly transfectable variant from mouse mastocytoma P815. *Somatic Cell Mol. Genet.* **11:** 467.

Voss, S.D., U. Schlokat, and P. Gruss. 1986. The role of enhancers in the regulation of cell-type-specific transcriptional control. *Trends Biochem. Sci.* **11:** 287.

Wahl, G.M., B. Robert de Saint Vincent, and M.L. DeRose. 1984. Effect of chromosomal position on amplification of transfected genes in animal cells. *Nature* **307:** 516.

Warden, D. and H.V. Thorne. 1968. Infectivity of polyoma virus DNA for mouse embryo cells in presence of diethylaminoethyl-dextran. *J. Gen. Virol.* **3:** 371.

Wieringa, B., F. Meyer, J. Reiser, and C. Weissmann. 1983. Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β-globin gene. *Nature* **301:** 38.

Wigler, M., S. Silverstein, L.-S. Lee, A. Pellicer, Y.-C. Cheng, and R. Axel. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* **11:** 223.

Wigler, M., R. Sweet, G.K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel. 1979. Transformation of mammalian cells with genes from procaryotes and eucaryotes. *Cell* **16:** 777.

Wong, G.G., J.S. Witek, P.A. Temple, K.M. Wilkens, A.C. Leary, D.P. Luxenberg, S.S. Jones, E.L. Brown, R.M. Kay, E.C. Orr, C. Shoemaker, D.W. Golde, R.J. Kaufman, R.M. Hewick, E.A. Wang, and S.C. Clark. 1985. Human GM-CSF: Molecular cloning of the complementary DNA and purification of the natural and recombinant proteins. *Science* **228:** 810.

Wong, T.-K. and E. Neumann. 1982. Electric field mediated gene transfer. *Biochem. Biophys. Res. Commun.* **107:** 584.

Wood, W.I., D.J. Capon, C.C. Simonsen, D.L. Eaton, J. Gitschier, B. Keyt, P.H. Seeburg, D.H. Smith, P. Hollingshead, K.L. Wion, E. Delwart, E.G.D. Tuddenham, G.A. Vehar, and R.M. Lawn. 1984. Expression of active human factor VIII from recombinant DNA clones. *Nature* **312:** 330.

Yamasaki, K., T. Taga, Y. Hirata, H. Yawata, Y. Kawanishi, B. Seed, T. Taniguchi, T. Hirano, and T. Kishimoto. 1988. Cloning and expression of the human interleukin-6 (BSF-2/IFNβ 2) receptor. *Science* **241:** 825.

Yang, Y.-C., A.B. Ciarletta, P.A. Temple, M.P. Chung, S. Kovacic, J.S. Witek-Giannotti, A.C. Leary, R. Kriz, R.E. Donahue, G.G. Wong, and S.C. Clark. 1986. Human IL-3 (multi-CSF): Identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. *Cell* **47:** 3.

Yates, J.L., N. Warren, and B. Sugden. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature* **313:** 812.

Yeung, C.-Y., D.E. Ingolia, D.B. Roth, C. Shoemaker, M.R. Al-Ubaidi, J.-Y. Yen, C. Ching, C. Bobonis, R.J. Kaufman, and R.E. Kellems. 1985. Identification of functional murine adenosine deaminase cDNA clones by complementation in *Escherichia coli*. *J. Biol. Chem.* **260:** 10299.

Zimmermann, U. 1982. Electric field-mediated fusion and related electrical phenomena. *Biochim. Biophys. Acta* **694:** 227.

Zinn, K., D. DiMaio, and T. Maniatis. 1983. Identification of two distinct regulatory regions adjacent to the human β-interferon gene. *Cell* **34:** 865.

17
Expression of Cloned Genes in Escherichia coli

Methods for expressing large amounts of protein from a cloned gene introduced into *Escherichia coli* have proven invaluable in the purification, localization, and functional analysis of proteins. For example, fusion proteins consisting of amino-terminal peptides encoded by a portion of the *E. coli lacZ* or *trpE* gene linked to eukaryotic proteins have been used to prepare polyclonal and monoclonal antibodies against these eukaryotic proteins. These antibodies have been used (1) to purify proteins by immunoaffinity chromatography; (2) in diagnostic assays to quantitate the levels of protein; and (3) to localize the proteins in organisms, tissues, and individual cells by immunofluorescence. Intact native proteins have also been produced in *E. coli* in large amounts for functional studies. For example, both prokaryotic and eukaryotic DNA-binding proteins produced using *E. coli* expression vectors have been used to study the role of these proteins in gene expression. In addition, proteins encoded by viral and cellular oncogenes have been produced in bacteria and then shown to be biologically active after microinjection into mammalian cells in culture. Finally, protein engineering studies are based on methods for introducing mutations into genes that encode proteins of interest and for producing the mutant proteins in large amounts in bacteria.

In this chapter, we describe methods and plasmid vectors for producing fusion proteins and intact native proteins in bacteria. Fusion proteins can be made in large amounts, are easy to purify, and can be used to elicit an antibody response. (Procedures for raising antibodies against fusion proteins in rabbits are given in Chapter 18, pages 18.5–18.6.) Native proteins can be produced in bacteria by placing a strong, regulated promoter and an efficient ribosome-binding site upstream of the cloned gene. If low levels of protein are produced, additional steps may be taken to increase protein production; if high levels of protein are produced, purification is relatively easy. Often, proteins expressed at high levels are found in insoluble inclusion bodies. Methods for extracting proteins from these aggregates are described on pages 17.37–17.38.

Many of the procedures utilized in this chapter (e.g., ligation, transformation, screening, and plasmid minipreps) are described in detail in Chapter 1. Readers who are not familiar with these procedures should refer to the relevant sections there.

Production of Fusion Proteins

If antibodies to the protein encoded by the cloned DNA are not available, it is often useful to produce a fusion protein for use as an antigen. Gene fusions in which DNA encoding part of the cloned gene is inserted near the 3′ terminus of the *lacZ* gene have been particularly useful for the following reasons:

- The fusion protein is usually made at high levels because transcription and translation initiation are directed by normal *E. coli* sequences.
- Fusion of foreign sequences to *E. coli* genes often results in products that are more stable than the native foreign proteins.
- The fusion protein is larger than most *E. coli* proteins and is therefore easy to identify in a protein gel. The fusion protein band can be cut out of the gel, lyophilized, ground into a powder, and used as an antigen.

The procedures presented below are designed to allow production of large amounts of a readily isolatable antigen.

Vector Systems for the Expression of lacZ Fusion Genes

Several vector systems have been developed for the expression of *lacZ* fusion genes. One such system, the pUR series of vectors (Rüther and Müller-Hill 1983), has cloning sites for *Bam*HI, *Sal*I, *Pst*I, *Xba*I, *Hin*dIII, and *Cla*I in all three reading frames at the 3′ terminus of the *lacZ* gene (see Figure 17.1 for a description of which vectors contain which cloning sites). By the appropriate choice of vector and restriction site, it should be possible to construct an in-frame fusion for virtually any cloned gene. In some cases, the fusions can be made by removing protruding termini or filling recessed termini prior to ligation.

A similar vector system developed for the production of fusion proteins uses bacterial expression vectors pEX1–3 (Stanley and Luzio 1984) in which the bacteriophage λ p_R promoter directs the expression of large amounts of a Cro–β-galactosidase fusion protein (see Figure 17.2). The bacteriophage λ p_R promoter is regulated and induced in the same manner as the bacteriophage λ p_L promoter (see page 17.11). The vectors pEX1–3 each contain a polycloning site in a different translational reading frame with *Eco*RI, *Sma*I, *Bam*HI, *Sal*I, and *Pst*I sites located in the 3′ end of the *lacZ* gene. Translation termination codons and transcription stop signals have been placed downstream from the polycloning site.

An alternative approach to the production of fusion proteins employs open reading frame (ORF) vectors (Gray et al. 1982; Weinstock et al. 1983) (see Figure 17.3). The ORF vectors are designed so that *lacZ* is not in-frame with sequences encoding the amino terminus of the protein and they therefore express very low levels of functional β-galactosidase. Insertion of a DNA fragment that contains an open reading frame and that places the *lacZ* sequences in-frame results in a plasmid expressing a protein that has β-galactosidase activity. After cDNA fragments are inserted into the ORF vector, *E. coli* strain LG90 (in which *lacZ* has been deleted) is transformed and plated on MacConkey lactose indicator plates. Cells expressing high levels of β-galactosidase are screened for the proteins synthesized.

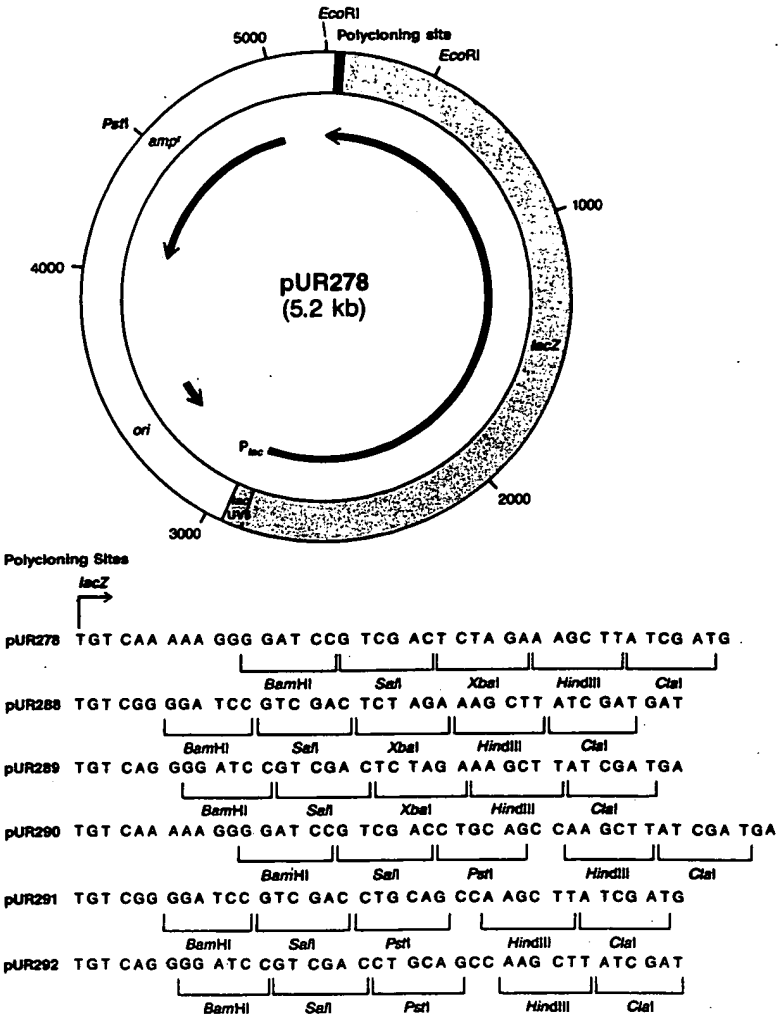


FIGURE 17.1
Vectors pUR278, pUR288, pUR289, pUR290, pUR291, and pUR292. These *lacZ* fusion vectors have the cloning sites *Bam*HI, *Sal*I, *Pst*I, *Xba*I, *Hin*dIII, and *Cla*I in all three reading frames at the 3′ terminus of the *lacZ* gene. Insertion of cDNA sequences in the appropriate cloning sites allows the production of a fusion protein of active β-galactosidase and the peptide encoded by the cDNA. pUR278, pUR288, and pUR289 plasmids contain polycloning sites that allow fusions in each of the three reading frames, and they contain a *Pst*I site in the amp^r gene. pUR290, pUR291, and pUR292 plasmids contain a unique *Pst*I site within the polycloning site because the *Pst*I site in amp^r has been destroyed. These plasmids are useful if the cDNA of interest was cloned into a *Pst*I site using the GC-tailing method.

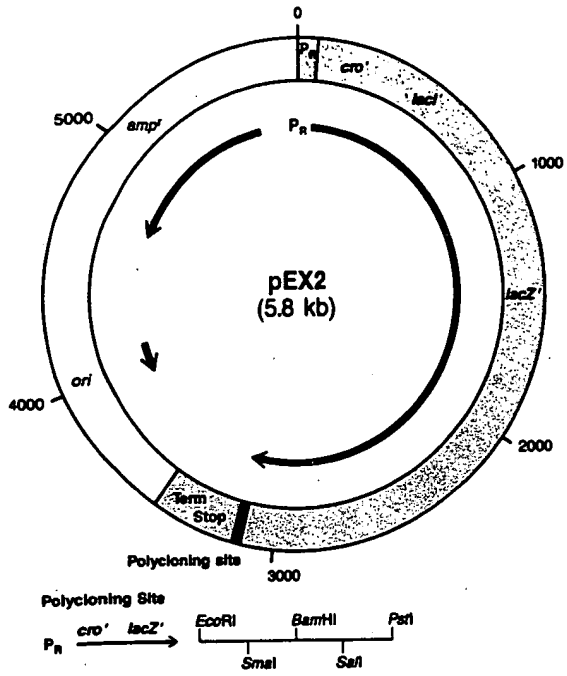


FIGURE 17.2
pEX2, a plasmid about 5.8 kb in length, is designed for expression of cDNA fused at the 3′ terminus of the *lacZ* gene. The amino-terminal part of the *lacZ* gene has been replaced with some sequences from the bacteriophage λ *cro* gene and the *E. coli lac*I gene. The bacteriophage λ p_R promoter is used, which allows expression of the fusion protein to be regulated by the bacteriophage λ c*Its*857 repressor. A polycloning site is present at the 3′ terminus of the *lacZ* gene, followed by the translation stop codons (Stop) and a transcription terminator (Term) from the bacteriophage fd (Stanley and Luzio 1984).

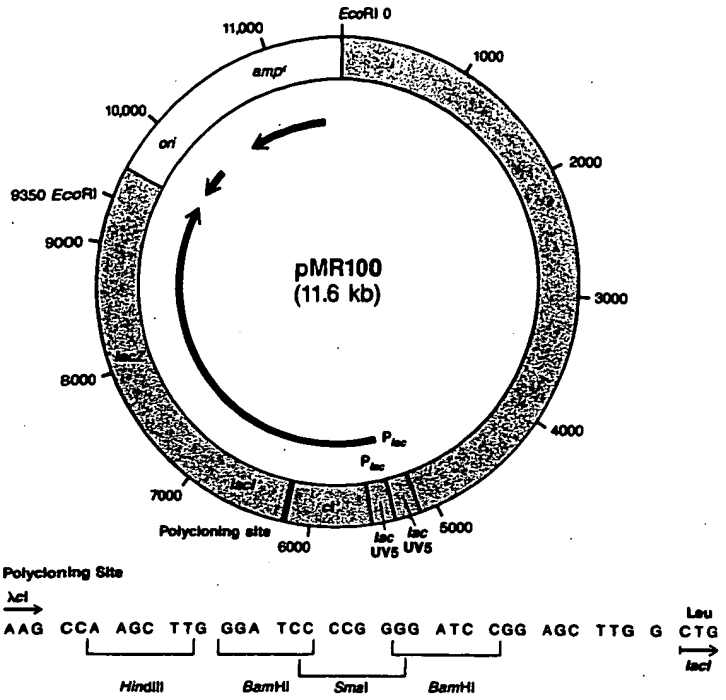


FIGURE 17.3

pMR100 is an ORF vector that carries a *lac* promoter, which directs the transcription of part of the *c*I gene of bacteriophage λ, followed by a polycloning site and an out-of-frame *lacZ* gene deleted at its 5′ terminus. When foreign DNA is inserted into the polycloning site such that the frameshift is corrected, a fusion protein is made that has the amino terminus of the bacteriophage λ *c*I gene, a middle segment of foreign protein sequence, and β-galactosidase at the carboxyl terminus. Cells making a fusion protein will express a high level of β-galactosidase, which can be detected by plating on MacConkey lactose plates (Gray et al. 1982).

1. Ligate the appropriate pUR (or pEX or pMR100) vector (see pages 17.4–17.7) and cDNA fragments to create an in-frame fusion.

2. Transform *E. coli* K12 71/18 or JM103 for pUR vectors (M5219 for pEX vectors or LG90 for pMR100) and plate on LB medium containing ampicillin (100 μg/ml). Incubate overnight at 37°C (or at 30°C if a pEX vector was used).

   MacConkey lactose indicator plates should be used for pMR100.

3. Test individual colonies for the presence of the desired insert by plasmid minipreps.

   If most of the colonies can be assumed to contain a cDNA (because directional cloning or a dephosphorylated vector was used in step 1), they can be screened for protein production in parallel (see step 4b). If not, clones that contain a cDNA, as determined by plasmid minipreps, can be screened for protein expression later.

   cDNA inserts into a pMR100 plasmid can be detected readily as red colonies on the MacConkey lactose indicator plates.

4. Screen colonies for the fusion protein.

   a. Grow small cultures from 5–10 colonies in LB medium containing ampicillin (100 μg/ml). Incubate overnight at 37°C (or at 30°C for pEX).

   b. Inoculate 5 ml of LB medium containing ampicillin (100 μg/ml) with 50 μl of each overnight culture. Incubate for 2 hours at 37°C (or at 30°C for pEX) with aeration. Remove 1 ml of uninduced culture, place it in a microfuge tube, and process as described in steps d and e. If screening for protein production is being done in parallel, prepare plasmid minipreps from 1-ml aliquots of the overnight cultures.

   c. Induce each culture as follows: For pUR or pMR100 vectors, add isopropylthio-β-D-galactoside (IPTG) to a final concentration of 1 mM and continue incubation at 37°C with aeration. For pEX vectors, transfer the culture to 40°C and continue incubating with aeration.

      For preparation of a stock solution of IPTG, see Appendix B.

   d. At various time points during the incubation (i.e., 1, 2, 3, and 4 hours), transfer 1 ml of each culture to a microfuge tube, and centrifuge at 12,000*g* for 1 minute at room temperature in a microfuge. Remove the supernatant by aspiration.

      The kinetics of induction varies with different proteins, so it is necessary to determine the time at which the maximum amount of product is produced.

   e. Resuspend each pellet in 100 μl of 1 × SDS gel-loading buffer, heat to 100°C for 3 minutes, and then centrifuge at 12,000*g* for 1 minute at room temperature. Load 15 μl of each suspension on a 6% SDS-polyacrylamide gel (see Chapter 18, pages 18.47–18.54). Use suspensions of cells containing the vector alone as a control. (For pEX and

ORF vectors, also use β-galactosidase [Sigma] as a control.) The fusion protein should appear as a novel band migrating more slowly than the intense β-galactosidase band in the control. It is not uncommon for a protein the size of β-galactosidase to be present along with the fusion protein.

1× SDS gel-loading buffer

50 mM Tris·Cl (pH 6.8)
100 mM dithiothreitol
2% SDS (electrophoresis grade)
0.1% bromophenol blue
10% glycerol

1× SDS gel-loading buffer lacking dithiothreitol can be stored at room temperature. Dithiothreitol should then be added, just before the buffer is used, from a 1 M stock (see Appendix B).

Preparation of Fusion Proteins for Antibody Production

The fusion protein can be prepared for injection in a number of ways, including urea extraction of insoluble aggregates (Schoner et al. 1985) (see page 17.40), aminophenylthiogalactoside affinity chromatography (Germino and Bastia 1984), preparative SDS-polyacrylamide gel electrophoresis, or any combination of these. One simple approach is to scale up the screening procedure.

1. Based on the optimal induction time determined in the screening procedure (step 4, page 17.8), grow and induce a scaled-up version (e.g., 200 ml) of the culture as described on page 17.8, step 4a–c. Transfer the culture to centrifuge tubes, and centrifuge at 5000 rpm for 15 minutes at 4°C. Remove the supernatant, and resuspend the pellet in a volume of 1× SDS gel-loading buffer equivalent to one tenth of the original volume of the cell culture.

2. Using a 5% SDS-polyacrylamide gel with a single large well, load the maximum amount of protein that still forms a reasonably tight band. (For preparation of SDS-polyacrylamide gels, see Chapter 18, pages 18.47–18.54.) Be certain from control experiments (page 17.8, step 4e) that the band of interest is clearly identifiable.

3. Stain the gel by soaking it in cold 0.25 M KCl (Hager and Burgess 1980) for 10–20 minutes with gentle agitation.

4. Excise the band from the gel, lyophilize for about 2 days, and then grind it into a fine powder. This powder is then used for injection into rabbits as described in Chapter 18, pages 18.5–18.6.

Intact native proteins can be made in *E. coli* by providing a strong, regulated promoter and an efficient ribosome-binding site. To express a prokaryotic gene that has a strong ribosome-binding site, only a promoter must be supplied (see pages 17.11–17.16). To express a eukaryotic gene (or a prokaryotic gene with a weak ribosome-binding site), both a promoter and a ribosome-binding site must be provided (see pages 17.17–17.25). Levels of expression may vary from less than 1% to more than 30% of total cell protein. Measures to improve expression are discussed on page 17.36.

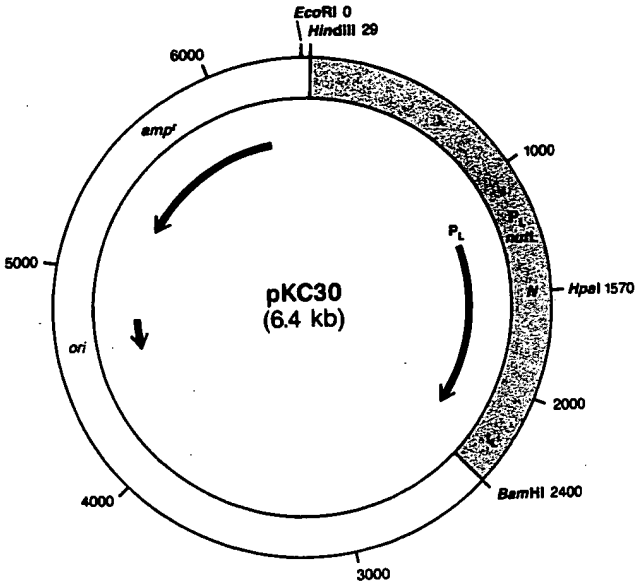


FIGURE 17.4
pKC30, a plasmid approximately 6.4 kb in length, carries the p_L promoter of bacteriophage λ and a *Hpa*I recognition site located 321 nucleotides downstream from the p_L transcriptional start site. The plasmid is a derivative of pBR322 and contains a *Hin*dIII-*Bam*HI fragment (gray box) derived from bacteriophage λ inserted between the *Hin*dIII and *Bam*HI sites within the tetracycline resistance gene (*tet*^{r}). The insertion contains the promoter signal, p_L, a site recognized by the *N* gene product (*nut*L), the *N* gene itself, and the strong *rho*-dependent transcription-termination signal t_L. The *Hpa*I recognition site lies within the coding region of the *N* gene. Sequences inserted into the *Hpa*I site can be regulated by introducing the recombinant plasmid into a temperature-sensitive bacteriophage λ lysogen (c*Its*857). The cells are grown to mid-log phase at 30°C and then shifted to 40°C to inactivate the *c*I gene product and to turn on the p_L promoter. This vector has been used to express the bacteriophage λ *c*II protein at a level such that the protein comprises 4% of the total protein of the cell (Shimatake and Rosenberg 1981).

EXPRESSION OF PROKARYOTIC GENES: PROMOTERS

The first step in expressing eukaryotic proteins in bacteria is to choose an expression vector that carries a strong, regulated prokaryotic promoter. Here we describe the use of expression vectors that contain a bacteriophage λ p_L promoter, a hybrid *trp-lac* promoter, or a bacteriophage T7 promoter.

The Bacteriophage λ p_L Promoter

The bacteriophage λ p_L promoter is regulated by a temperature-sensitive repressor, *cIts857*, which represses p_L transcription at low temperatures but not at elevated temperatures. *E. coli* strain M5219 contains a defective bacteriophage λ prophage, which encodes the *cIts857* repressor, and the bacteriophage λ N protein, an antagonist of transcription termination. This system is particularly suitable if the product of the gene to be expressed is toxic to *E. coli*, since *cIts857* strongly represses transcription. In addition, the antitermination function of the *N* gene may allow RNA polymerase to read through potential termination sites within the gene. One drawback of this system is that the temperature shift (step 5) induces not only the p_L promoter, but also the heat-shock genes, some of which encode proteases (Buell et al. 1985). This problem can be alleviated by using a bacteriophage λ cI^+ lysogen and inducing with mitomycin C or nalidixic acid (Shatzman and Rosenberg 1987). pKC30 (see Figure 17.4) (Shimatake and Rosenberg 1981) is one of many p_L vectors available.

To provide a p_L promoter to direct transcription of a cloned gene that has an *E. coli* ribosome-binding site:

1. Digest pKC30 with *Hpa*I.
2. Digest the cloned DNA with appropriate restriction enzymes at a position 5′ of the initiation codon and at a site 3′ of the terminus of the cloned gene.
3. Insert the cDNA into the plasmid, ligate the DNA, and transform *E. coli* strain M5219. Plate transformants on LB medium containing ampicillin (100 μg/ml) and incubate overnight at 30°C.
4. Screen transformants for the presence of the desired insert by colony hybridization and/or by plasmid minipreps and restriction enzyme analysis.
5. To obtain high levels of transcription of the cloned gene, grow strain M5219 containing the expression plasmid to mid-log phase at 30°C, and then shift the temperature of the culture to 40°C. Continue to incubate for several hours at 40°C. Remove small aliquots at various times and analyze them by one of the methods discussed on pages 17.34–17.35. The kinetics of induction varies with different proteins, so it is necessary to determine the time at which the maximum amount of product is present.

   Although 42–45°C is used elsewhere in this manual to inactivate bacteriophage λ *cIts857*, 40°C is used here in order to reduce the induction of heat-shock proteins and to allow the cells to continue growth.

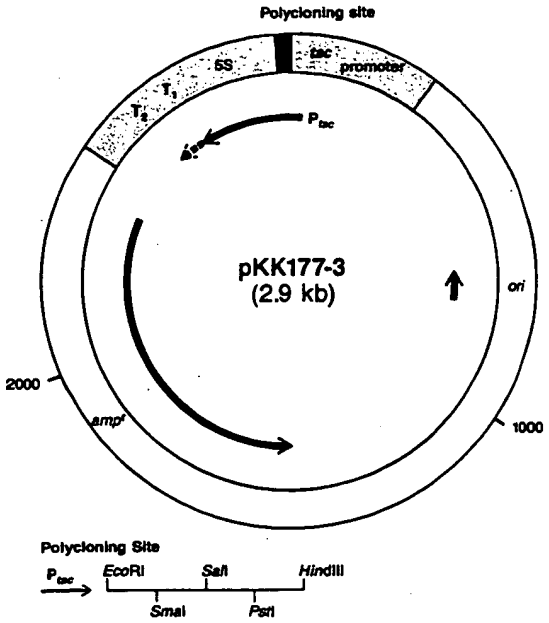


FIGURE 17.5
pKK177-3 is a *tac* vector containing multiple sites downstream from the *tac* promoter into which a gene can be cloned. Downstream from these sites is *rrn*B, which contains an *E. coli* 5S gene and the T_1 and T_2 terminators (Amann and Brosius 1985).

The trp-lac Promoter

Another promoter that has been used successfully to produce large amounts of proteins in *E. coli* is the *tac* promoter, a hybrid *trp-lac* promoter that is regulated by *lac* repressor (Amann et al. 1983; de Boer et al. 1983). Transcription is repressed in *E. coli* strains such as RB791, a *lacI*^{q} strain that makes high levels of *lac* repressor. If the protein to be expressed is toxic to *E. coli*, then a *lacI*^{q} gene should be cloned into the expression plasmid to make higher levels of *lac* repressor. One factor to consider in choosing this vector system is that the *tac* promoter is induced by adding isopropylthio-β-D-galactoside (IPTG), a relatively expensive compound. A useful *tac* promoter expression plasmid, pKK177-3 (Amann and Brosius 1985), is illustrated in Figure 17.5.

The steps used in cloning and expression in *tac* vectors (or *lac* vectors) are similar to those for other promoter systems and are described below.

1. Clone the cDNA to be expressed, along with its ribosome-binding site, into the polycloning site of pKK177-3.

2. Transform *E. coli* strain RB791, and screen transformants by colony hybridization and/or plasmid minipreps and restriction enzyme analysis.

3. To obtain high levels of transcription, grow cells to mid-log phase at 37°C, add IPTG to a final concentration of 1 mM, continue growth, and monitor the amount of protein made at various times after induction using one of the methods discussed on pages 17.34–17.35.

   For preparation of a stock solution of IPTG, see Appendix B.

Note

Hasan and Szybalski (1987) have designed *tac* vectors in which expression of a cloned gene is controlled by promoter inversion in vivo. The promoter in the vector is directed away from the gene to be expressed so that transcription of the gene is kept extremely low until production is desired. The promoter is flanked by *att*P and *att*B and can be inverted efficiently by the bacteriophage λ Int protein. These vectors also utilize bacteriophage λ *N*-mediated antitermination as described on page 17.11 for pKC30.

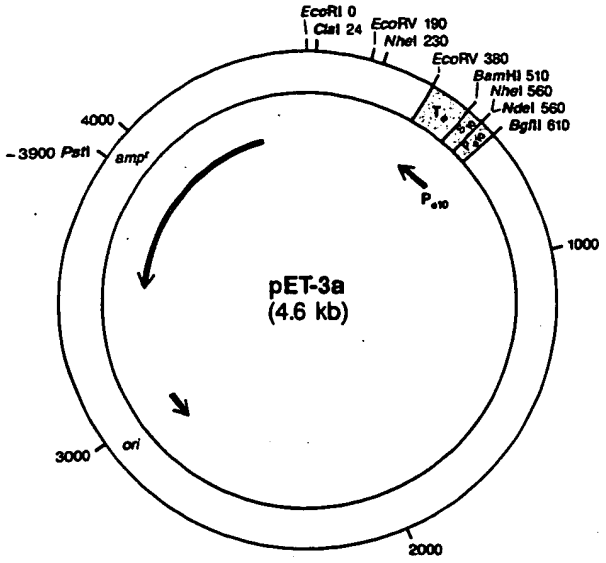


FIGURE 17.6
pET-3 carries the bacteriophage T7 ϕ10 promoter ($P_{\phi 10}$) and ϕ terminator (T_{ϕ}). The terminator may make the transcripts more resistant to exonucleolytic degradation (Studier and Moffatt 1986). pET-3a is a derivative of pET-3 into which the translation start (S_{10}) of bacteriophage T7 ϕ10 (the major capsid protein of bacteriophage T7) with a *Bam*HI site at codon 11 has been inserted. The *Nde*I site (CATATG) is located at the translation start site and can be used to construct a plasmid that directs the expression of native proteins.

The Bacteriophage T7 Promoter

A novel expression system has been developed by Studier and Moffatt (1986) and Tabor and Richardson (1985) using a bacteriophage T7 RNA polymerase/promoter system. This system is designed for the exclusive expression of the cloned gene. Bacteriophage T7 RNA polymerase recognizes solely bacteriophage T7 promoters, will transcribe around a plasmid several times, and may transcribe sequences that are not efficiently transcribed by *E. coli* RNA polymerase. This system allows high levels of expression of some genes that are not expressed efficiently in other systems.

Two components are required for the bacteriophage T7 expression system:

- The first component is bacteriophage T7 RNA polymerase. This polymerase is the product of bacteriophage T7 gene *1* and can be provided on an infecting bacteriophage λ vector or produced from a gene copy inserted into the *E. coli* chromosome (Tabor and Richardson 1985; Studier and Moffatt 1986). If expression of the cloned gene product is toxic, then the level of bacteriophage T7 RNA polymerase must be kept low during cell growth. One way to accomplish this is to use the lysogen BL21(DE3), in which bacteriophage T7 gene *1* is expressed from a *lac*UV5 promoter. In some cases, it is necessary to use cells in which no bacteriophage T7 RNA polymerase is present until expression is desired. This is accomplished by infecting the host cells (e.g., HMS174) that carry the expression plasmid with bacteriophage CE6 (λc*Its*857 *S*am7) carrying bacteriophage T7 gene *1* (Studier and Moffatt 1986).
- The second component is a plasmid vector with a bacteriophage T7 promoter upstream of the gene to be expressed. pET-3 is a derivative of pBR322 that carries the bacteriophage T7 gene *10* promoter ($P_{\phi 10}$), a *Bam*HI cloning site, and the bacteriophage T7 transcription terminator (T_{ϕ}) (Rosenberg et al. 1987). Derivatives of pET-3 (e.g., pET-3a; see Figure 17.6) have been constructed to provide a bacteriophage T7 gene *10* translation start (S_{10}) through codon 11. DNA can be inserted in each of the three reading frames at this codon to express fusion proteins. The pET-3 derivatives can also be used for production of intact native proteins as described on pages 17.17–17.24 by fusing coding sequences at the *Nde*I site immediately preceding the ATG. Another bacteriophage T7 expression plasmid, pT7-1, has a polycloning site downstream from the bacteriophage T7 ϕ10 promoter (Tabor and Richardson 1985).

To express a gene in *E. coli* using bacteriophage T7 RNA polymerase:

1. Clone the gene of interest into a bacteriophage T7 promoter expression plasmid. Identify the correct plasmids in a standard *E. coli* strain by miniprep analysis.

2. Transform *E. coli* strain BL21(DE3), and select for ampicillin-resistant transformants. BL21(DE3) is a lysogen bearing the bacteriophage T7 polymerase gene under the control of the *lac*UV5 promoter. (See precautions recommended by Studier and Moffatt [1986].)

3. Inoculate NZCYM medium with one colony, and incubate overnight at 37°C to obtain a saturated culture.

4. Induce the culture and determine the amount of protein produced as follows:

   a. Inoculate 5 ml of NZCYM medium containing ampicillin (100 μg/ml) with 50 μl of saturated culture. Incubate the culture for 2 hours at 37°C.

   b. Remove 1 ml of the uninduced culture and place in a microfuge tube. Process as described in steps c and d below. Induce the remaining culture by adding isopropylthio-β-D-galactoside (IPTG) to a final concentration of 1 mM.

      For preparation of a stock solution of IPTG, see Appendix B.

   c. Remove 1-ml aliquots of the induced culture at 0.5, 1, 2, and 3 hours after induction. Immediately centrifuge these samples at 12,000*g* for 1 minute at room temperature in a microfuge. Remove the supernatants by aspiration.

   d. Resuspend each pellet in 100 μl of 1 × SDS gel-loading buffer, heat to 100°C for 3 minutes, and store at 0°C until all of the samples are collected and ready to load on a gel.

      1 × SDS gel-loading buffer

      50 mM Tris·Cl (pH 6.8)
      100 mM dithiothreitol
      2% SDS (electrophoresis grade)
      0.1% bromophenol blue
      10% glycerol

      1× SDS gel-loading buffer lacking dithiothreitol can be stored at room temperature. Dithiothreitol should then be added, just before the buffer is used, from a 1 M stock (see Appendix B).

   e. Thaw the samples, and then centrifuge them at 12,000*g* for 1 minute at room temperature. Load 15 μl of each suspension on an SDS-polyacrylamide gel of the appropriate concentration (see Chapter 18, pages 18.47–18.54). Use as a control a suspension of cells containing the vector alone.

5. Stain the gel with Coomassie Brilliant Blue or silver stain, or carry out a western blot (if antibody already exists), to identify induced protein (see Chapter 18, pages 18.55–18.57 or 18.60–18.75, respectively).

Note

For expression of some proteins, it is advisable to increase the ampicillin concentration to levels as high as 200 μg/ml to select for cells that retain the expression plasmid.

EXPRESSION OF EUKARYOTIC GENES: PROM TERS AND RIBOSOME-BINDING SITES

In addition to the bacterial promoter, the second major factor required to express cloned DNA in *E. coli* is an efficient ribosome-binding site. In *E. coli*, the ribosome-binding site includes an initiation codon (ATG) and a sequence 3–9 nucleotides in length located 3–11 nucleotides upstream of the initiation codon (Shine and Dalgarno 1975; Steitz 1979). This sequence, which is called the Shine-Dalgarno (SD) sequence, is complementary to the 3′ terminus of *E. coli* 16S rRNA. Binding of the ribosome to mRNA is thought to be promoted by base pairing between the SD sequence in the mRNA and the sequence at the 3′ terminus of the 16S rRNA (Steitz 1979).

When expressing a prokaryotic gene, the ribosome-binding site of that gene may be sufficient. In this case, the protocols presented earlier, which place the gene downstream from a strong, regulated promoter and employ a restriction site 5′ of the ribosome-binding site, may allow high levels of expression. Eukaryotic genes and prokaryotic genes with weak ribosome-binding sites require that an efficient ribosome-binding site be provided. This can be accomplished in several ways, as described below.

Preparation of a DNA Fragment for Placement Adjacent to a Functional Ribosome-binding Site

To insert a cloned gene into a vector that carries a functional ribosome-binding site (an initiator ATG an appropriate distance from an SD sequence), the second codon of the cloned gene must be placed adjacent to the initiator ATG. A DNA fragment extending from the second codon to a restriction site 3′ of the cloned gene is made by one of the methods given below.

SYNTHESIS OF DNA ENCODING THE AMINO TERMINUS

1. Synthesize a double-stranded DNA adapter extending from the second codon of the gene (blunt end) to a restriction site within the gene (sticky end).

2. Ligate this adapter to a fragment of the gene extending from the restriction site used in step 1 to a site downstream from the end of the gene.

3. Depending on the choice of vector and availability of compatible restriction sites, it may not be necessary to create a blunt end at the downstream site or it may be possible to use a restriction enzyme that leaves a blunt end. If necessary, repair the DNA from step 2 with the Klenow fragment of *E. coli* DNA polymerase I to create a blunt end at the downstream site (see page 17.23, step 5a–c).

4. Isolate the DNA fragment containing the gene following gel electrophoresis to obtain a DNA fragment that can be inserted after the ATG of plasmid pAS1 (see Figure 17.8) or pKK 240-11 (see Figure 17.9).

PRIMER REPAIR

1. Synthesize a DNA primer whose 5′ terminus encodes the second amino acid of the protein to be expressed.

2. Phosphorylate the 5′ terminus of the primer with bacteriophage T4 polynucleotide kinase.

   a. Mix:

primer	100 pmoles
H_2O	42 μl
10 × bacteriophage T4 polynucleotide kinase buffer	5 μl
10 mM ATP	1 μl
bacteriophage T4 polynucleotide kinase (10 units)	1 μl

10 × Bacteriophage T4 polynucleotide kinase buffer

500 mM Tris · Cl (pH 7.6)
100 mM $MgCl_2$
50 mM dithiothreitol
1 mM spermidine HCl
1 mM EDTA (pH 8.0)

b. Incubate the reaction for 30 minutes at 37°C.

c. Heat the reaction for 5 minutes at 68°C to inactivate the bacteriophage T4 polynucleotide kinase. Place on dry ice and store at −20°C.

3. Hybridize the primer to template bacteriophage M13 single-stranded DNA carrying the cloned gene, and extend the primer with the Klenow fragment of *E. coli* DNA polymerase I.

a. Mix:

phosphorylated primer	1 μl
DNA template	1 μl
0.5 M NaCl	1 μl
H_2O	6 μl

b. Incubate the mixture for 5 minutes at 70°C to denature the DNAs.

c. Slowly cool to 18°C.

d. Add 10 μl of 2 × elongation mix and 2 μl of the Klenow fragment of *E. coli* DNA polymerase I. Incubate for 30 minutes at room temperature.

2 × Elongation mix

20 mM Tris · Cl (pH 7.5)
10 mM $MgCl_2$
10 mM dithiothreitol
a mixture of all four dNTPs, each at a concentration of 1 mM

4. Treat the mixture with nuclease S1 to remove single-stranded DNA.

a. Add 2 μl of 10 × nuclease-S1 buffer and 1 unit of nuclease S1.

10 × Nuclease-S1 buffer

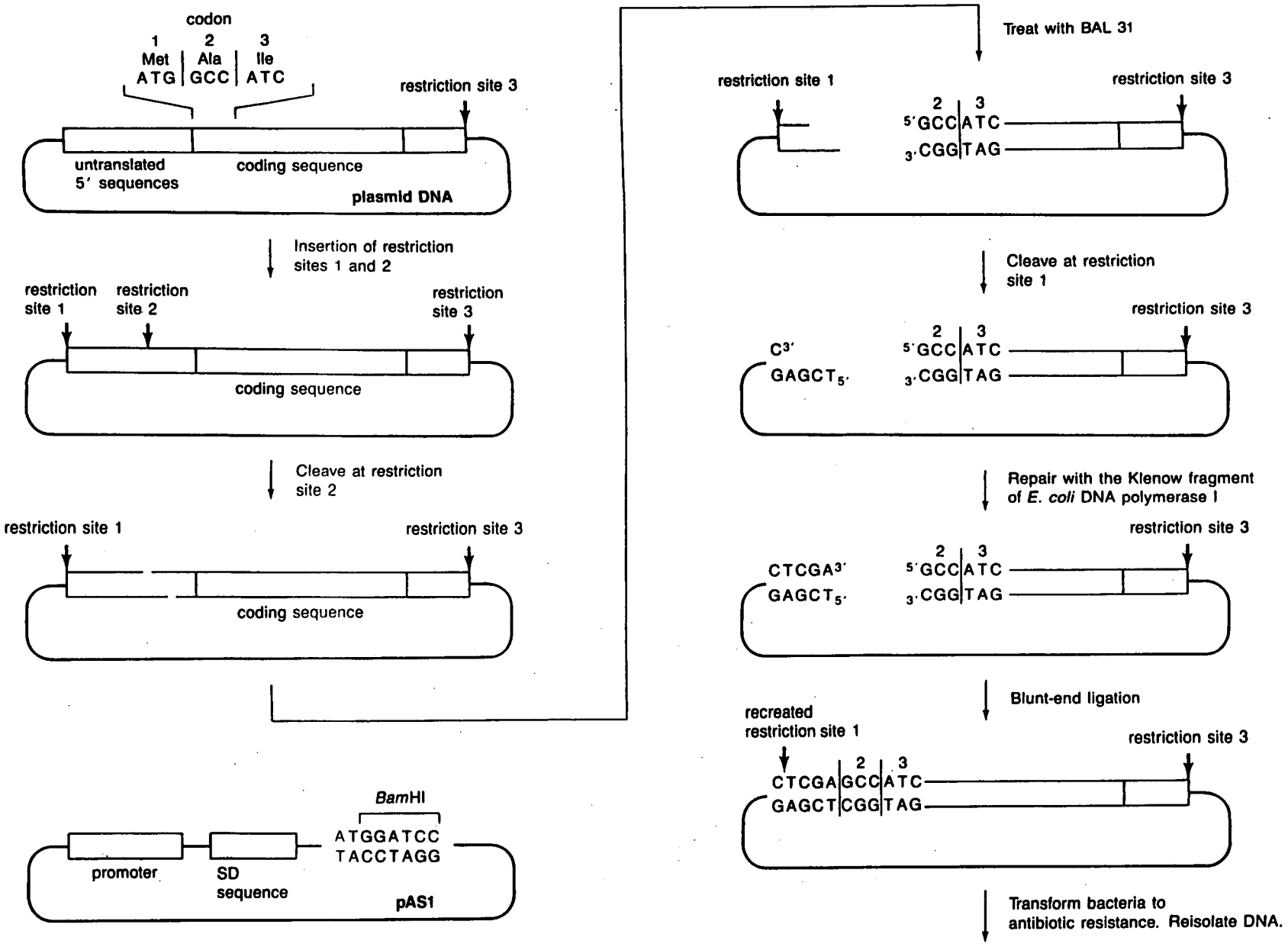
2 M NaCl
0.5 M sodium acetate (pH 4.5)
10 mM $ZnSO_4$
5% glycerol

b. Incubate the reaction for 30 minutes at 37°C.

c. Extract the reaction once with phenol:chloroform. Transfer the aqueous phase to a fresh microfuge tube, and precipitate the DNA with 2 volumes of ethanol at 0°C. Recover the DNA by centrifugation at 12,000*g* for 10 minutes at 0°C in a microfuge. Remove the supernatant by aspiration, and resuspend the DNA in the appropriate 1 × restriction enzyme buffer.

5. Digest the DNA with an enzyme that cleaves at a downstream site.

6. Continue with steps 3 and 4, page 17.18.


codon
1 2 3
Met | Ala | Ile
ATG | GCC | ATC
restriction site 3
untranslated 5′ sequences
coding sequence
plasmid DNA
Insertion of restriction sites 1 and 2
restriction site 1
restriction site 2
restriction site 3
coding sequence
Cleave at restriction site 2
restriction site 1
restriction site 3
coding sequence
BamHI
ATGGATCC
TACCTAGG
promoter
SD sequence
pAS1
Treat with BAL 31
restriction site 1
restriction site 3
2 3
5′GCC|ATC
3′CGG|TAG
Cleave at restriction site 1
restriction site 3
C3′
GAGCT5′
2 3
5′GCC|ATC
3′CGG|TAG
Repair with the Klenow fragment of E. coli DNA polymerase I
restriction site 3
CTCGA3′
GAGCT5′
2 3
5′GCC|ATC
3′CGG|TAG
Blunt-end ligation
recreated restriction site 1
restriction site 3
2 3
CTCGA|GCC|ATC
GAGCT|CGG|TAG
Transform bacteria to antibiotic resistance. Reisolate DNA.

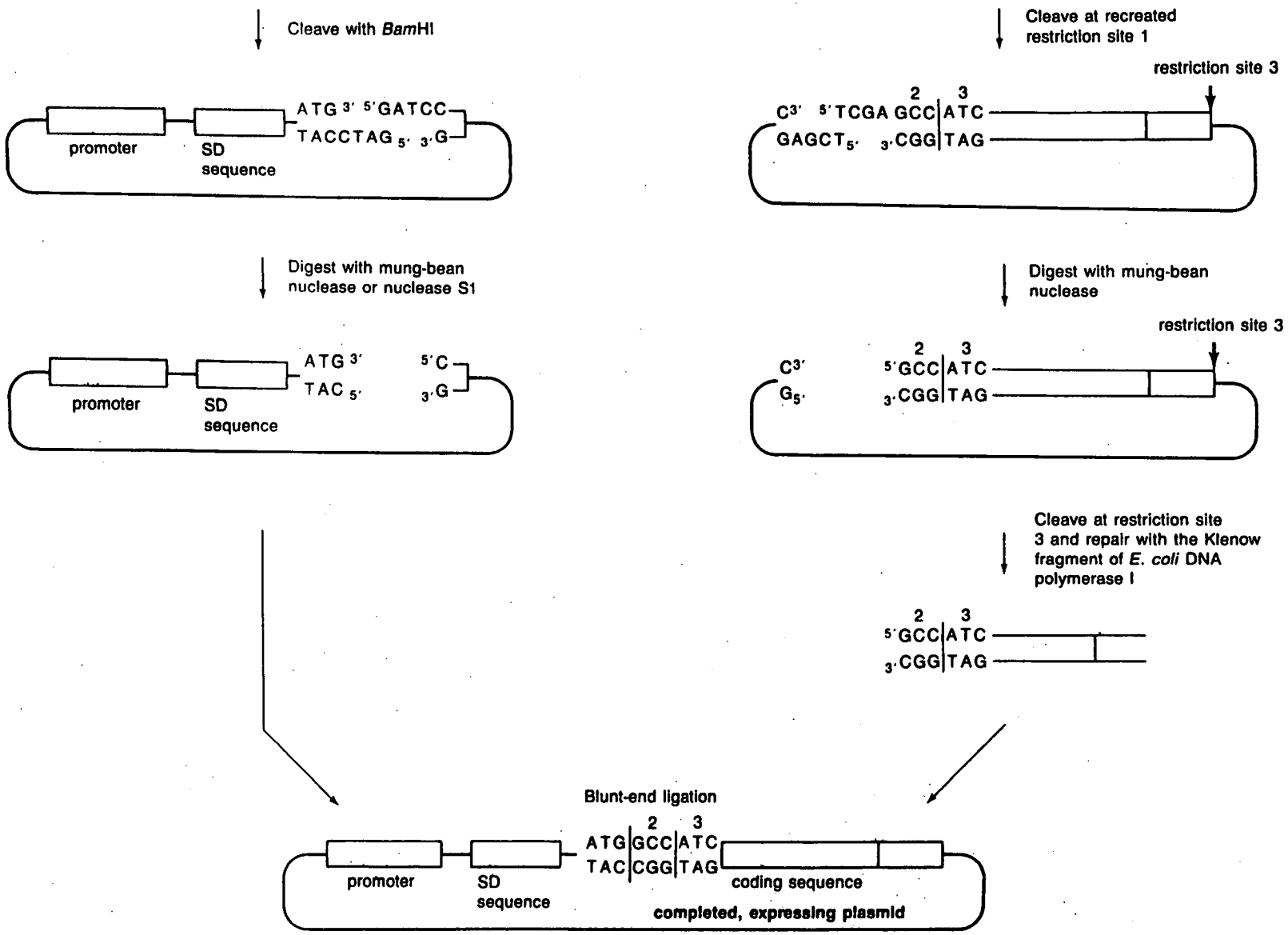


FIGURE 17.7
A method to create a blunt end immediately before a particular nucleotide in a segment of cloned DNA (see text for details).

Place a restriction site immediately preceding the second codon (Panayotatos and Truong 1981; Shatzman et al. 1983). This can be done as follows (see Figure 17.7, pages 17.20–17.21):

1. Insert two unique restriction sites upstream of the coding sequence to be expressed by inserting linkers or adapters or by cloning the gene into a plasmid that already has the two sites. Choose site 1 such that the sixth nucleotide is the first nucleotide of the second codon of the protein to be expressed and the enzyme leaves a 5′ overhang upon digestion. Site 2 must lie between site 1 and the start of the gene, and it must be closer to the start of the gene than to site 1.
2. Cleave 5 μg of the DNA at site 2 with the appropriate restriction enzyme.
3. Digest the DNA with nuclease BAL 31 to produce a population of molecules that randomly terminate close to the desired codon as follows:
   a. Precipitate the digested DNA with 2 volumes of ethanol at 0°C. Recover the DNA by centrifugation at 12,000*g* for 10 minutes at 0°C in a microfuge.
   b. Remove the supernatant by aspiration, and redissolve the DNA in 50 μl of bovine serum albumin (Fraction V; Sigma) (500 μg/ml in water).
   c. Add 50 μl of 2× BAL 31 buffer.

2× BAL 31 buffer
1.2 M NaCl
24 mM $CaCl_2$
24 mM $MgCl_2$
40 mM Tris · Cl (pH 8.0)
0.4 mM EDTA (pH 8.0)

   d. Incubate the mixture for 30 minutes at 30°C.
   e. Add 0.5 unit of BAL 31, which removes approximately 10 bp/minute/terminus.
   f. At five appropriate, closely spaced time points centered around the time needed to delete the bases between site 2 and the second codon of the gene, remove 10-μl samples and add 10 μl of 0.2 M EGTA (pH 8.0) to each sample. Store the samples on ice.
   g. Combine the samples, extract once with phenol:chloroform, and transfer the aqueous phase to a fresh microfuge tube.
   h. Precipitate the DNA with 2 volumes of ethanol at 0°C. Recover the DNA by centrifugation at 12,000*g* for 10 minutes at 0°C. Remove the supernatant, and wash the pellet with 70% ethanol. Redissolve the DNA in 20 μl of the appropriate 1× restriction enzyme buffer.
4. Cleave the DNA at site 1 with the appropriate restriction enzyme.

5. Repair the DNA with the Klenow fragment of *E. coli* DNA polymerase I to create a blunt end containing 5 of the 6 nucleotides of restriction site 1.

   a. Mix the DNA with:

10 × nick-translation buffer	2.5 μl
a mixture of all four dNTPs, each at a concentration of 2 mM	1 μl
Klenow fragment of *E. coli* DNA polymerase I	1 μl
H_2O to 25 μl	

10 × Nick-translation buffer

0.5 M Tris · Cl (pH 7.5)
0.1 M $MgSO_4$
1 mM dithiothreitol
500 μg/ml bovine serum albumin (Fraction V; Sigma) (optional)

Divide the stock solution into small aliquots and store them at −20°C.

   b. Incubate the reaction for 30 minutes at room temperature.

   c. Stop the reaction by adding 1 μl of 0.5 M EDTA (pH 8.0).

   d. Extract once with phenol:chloroform.

   e. Separate the DNA from unincorporated dNTPs by chromatography on, or centrifugation through, small columns of Sephadex G-50 (see Appendix E).

6. Recircularize the DNA by ligation, and use the plasmid to transform bacteria to antibiotic resistance.

7. Screen individual colonies by plasmid minipreps and restriction enzyme analysis for the presence of plasmids in which restriction site 1 has been regenerated. Such regeneration occurs when digestion with BAL 31 generates, by chance, a DNA molecule whose terminus carries the particular base pair required to complete site 1. Limited DNA sequence (see Chapter 13) may be required to identify the plasmid within the subpopulation in which the last nucleotide of site 1 is the first nucleotide of the second codon of the gene to be expressed.

8. Cleave 5 μg of the desired plasmid at site 1. Repeat step 3g–h, except resuspend the DNA in 10 μl of 10 mM Tris · Cl (pH 7.5). Digest with mung-bean nuclease to generate a blunt end directly preceding the first nucleotide of the second codon as follows:

   a. Mix the 10 μl of DNA with:

10 × mung-bean nuclease buffer	2 μl
mung-bean nuclease (1.5 units/μl)	2.5 μl
H_2O	5.5 μl

10 × Mung-bean nuclease buffer

300 mM sodium acetate (pH 4.5)
500 mM NaCl
10 mM $ZnCl_2$
50% glycerol

b. Incubate the reaction for 1 hour at 37°C.

c. Extract the reaction once with phenol:chloroform, and transfer the aqueous phase to a fresh microfuge tube.

d. Precipitate the DNA with 2 volumes of ethanol at 0°C. Recover the DNA by centrifugation at 12,000*g* for 10 minutes at 0°C in a microfuge. Remove the supernatant, and wash the pellet with 70% ethanol. Redissolve the DNA in 20 μl of the appropriate 1 × restriction enzyme buffer.

9. Cleave the DNA with a restriction enzyme that cuts downstream from the gene (site 3 in Figure 17.7).

10. Depending on the choice of vector and availability of compatible restriction sites, it may not be necessary to create a blunt end at the downstream site or it may be possible to use a restriction enzyme that leaves a blunt end. If necessary, repair the DNA from step 9 with the Klenow fragment of *E. coli* DNA polymerase I to create a blunt end at the downstream site (see step 5a–c).

11. Isolate the DNA fragment containing the gene following gel electrophoresis to obtain a DNA fragment that can be inserted after the ATG of plasmid pAS1 (see Figure 17.8) or pKK240-11 (see Figure 17.9).

Expression of a Gene from the Prepared Fragments

To express any gene:

1. Digest pAS1 (Figure 17.8) with *Bam*HI and remove protruding termini with mung-bean nuclease or nuclease S1 (see step 8, page 17.23, or step 4, page 17.19, respectively). pAS1 provides a bacteriophage λ promoter, p_L, and the ribosome-binding site of the bacteriophage λ *c*II gene. Alternatively, digest pKK240-11 (Figure 17.9) with *Nco*I and fill in recessed termini with the Klenow fragment of *E. coli* DNA polymerase I (see step 5, page 17.23). pKK240-11 provides a *tac* promoter and a *lacZ* ribosome-binding site.

2. If necessary, cleave the vector with a restriction enzyme that cuts at a downstream site compatible with the DNA fragment prepared by one of the methods described above (see pages 17.18–17.24).

   In most cases, a blunt-ended fragment will be inserted and this step will not be required. In a few cases, the gene fragment prepared as described on pages 17.18–17.24 will contain a compatible sticky end with a site within the vector.

3. Ligate the purified DNA fragment (see pages 17.18–17.24) to the vector prepared in steps 1 and 2, and transform the appropriate strain (see Chapter 1). For pKK240-11, use strain RB791; for pAS1, use M5219.

4. Screen transformants for colonies that contain the cloned gene by plasmid minipreps and restriction enzyme analysis. Sequence the junction between the ATG and the second codon (see Chapter 13). Alternatively, synthesize an oligonucleotide complementary to the correct sequence extending from the ribosome-binding site through the first part of the gene. Screen transformants by colony hybridization with the oligonucleotide that bridges the junction between the ATG and the second codon (see Chapter 11), and sequence this region in the positive clones (see Chapter 13).

5. To obtain high levels of transcription, follow the procedure described in step 5, page 17.11, or step 3, page 17.13, for induction of the p_L or *tac* promoter, respectively. Test the correct clones for protein production (see pages 17.34–17.35).

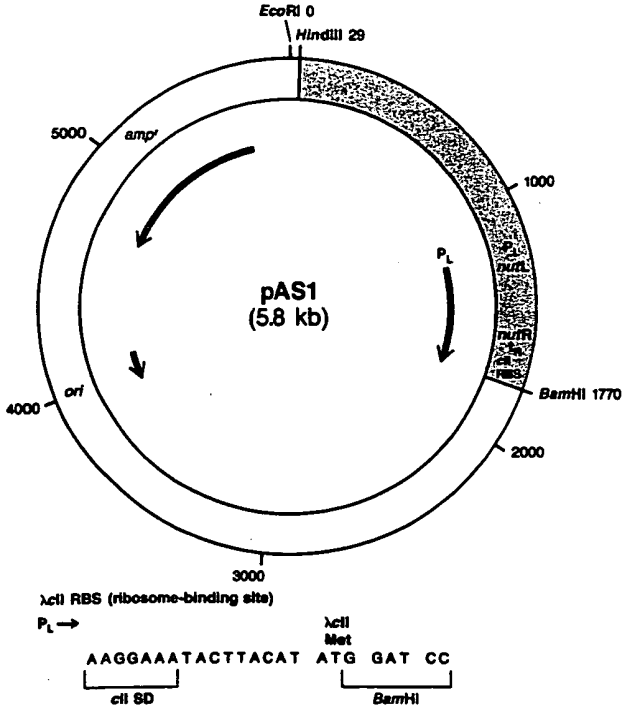


FIGURE 17.8

pAS1, a plasmid approximately 5.8 kb in length, carries the bacteriophage λ p_L promoter and a unique *Bam*HI site located at the ATG of the bacteriophage λ *c*II gene. This plasmid is a derivative of pKC30 (see Figure 17.4) into which the bacteriophage λ *c*II gene was inserted at the *Hpa*I site. The *c*II gene was then resected by exonuclease digestion until only the initiation codon ATG remained (the G of the ATG is the first nucleotide of a *Bam*HI site). To express a gene lacking an initiation codon, pAS1 is digested with *Bam*HI and then treated with mung-bean nuclease or nuclease S1 to remove the protruding, single-stranded termini. Ligation of this blunt-ended DNA to a blunt-ended DNA fragment that begins with the second codon of the gene to be expressed places that gene in-frame with the ATG. Genes inserted in this manner are regulated by introducing the recombinant plasmid into a temperature-sensitive bacteriophage λ lysogen (*cIts*857). The cells are grown to late log phase at 30°C and then shifted to 40°C to inactivate the repressor and to turn on the p_L promoter. The inserted gene can also be regulated by the action of the N protein at *nut*L and *nut*R to antiterminate at t_{R1} (Shatzman et al. 1983).

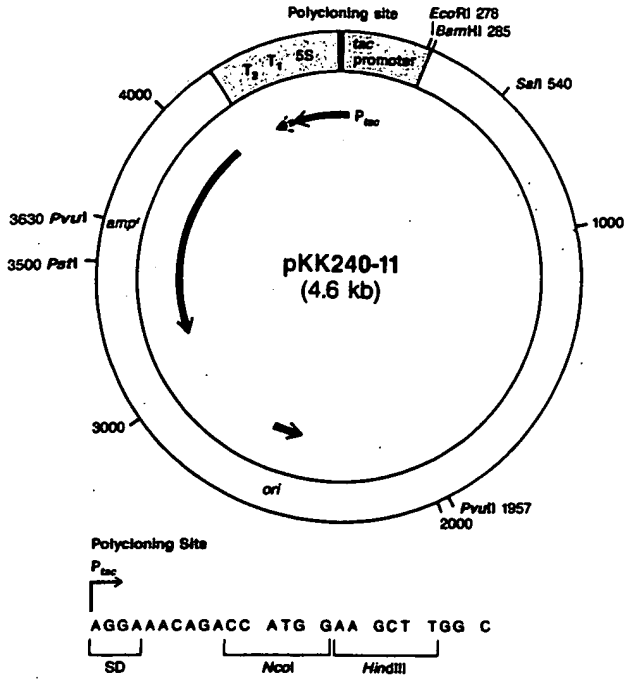


FIGURE 17.9

pKK240-11, a plasmid approximately 4.6 kb in length, carries a *tac* promoter, which is an *E. coli* hybrid promoter composed of *trp* and *lac* promoter sequences. A unique *Nco*I site, located at the ATG, can be filled in using the Klenow fragment of *E. coli* DNA polymerase I to form a blunt end to which the second codon of the gene to be expressed can be fused. Downstream from the polycloning site is *rrn*B, which contains an *E. coli* 5S gene and the T_1 and T_2 terminators (Amann and Brosius 1985).

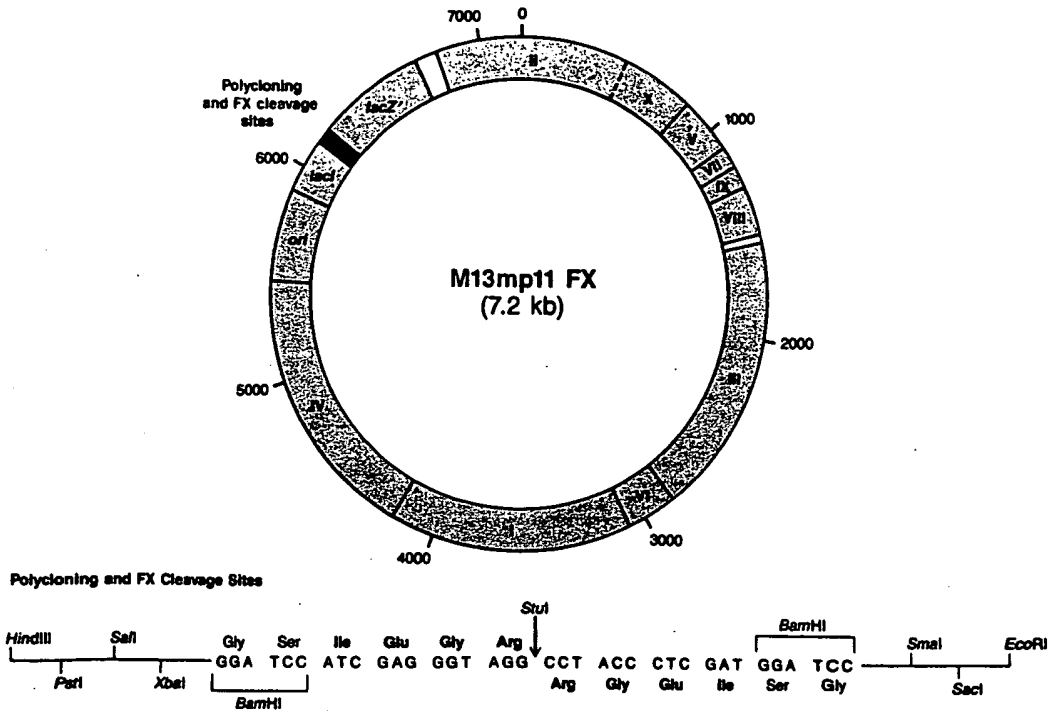


FIGURE 17.10

M13mp11 FX is an M13mp11 derivative that has the polycloning site oriented with *Eco*RI adjacent to *lacZ'*. DNA encoding the recognition sequence for blood coagulation factor X_a (Ile Glu Gly Arg [shown in bold]) is cloned into the *Bam*HI site. Any coding sequence with blunt ends can be cloned into the *Stu*I site to place the FX_a cleavage site adjacent to the desired protein sequence. The entire DNA sequence can then be inserted into an appropriate expression vector (e.g., pLcII) (Nagai and Thøgersen 1984).

ALTERNATIVE EXPRESSION SYSTEMS

In some cases, the methods described above do not result in the production of the desired protein. In particular, it may be important that the protein not have an amino-terminal methionine, which is coded for by the ATG in these vectors. The amino-terminal methionine is removed in *E. coli* to different extents for different proteins. Another common problem is that some proteins produced intracellularly are inactive, perhaps due to incorrect folding. Some foreign proteins are degraded by intracellular proteases. Two other expression systems may solve some of these problems.

Expression of a Cloned Gene as Part of a Fusion Protein That Can Be Cleaved by a Protease or Cyanogen Bromide

The cloned gene can be expressed as part of a fusion protein composed of a prokaryotic amino terminus connected to the protein of interest by a sequence that can be cleaved by a protease or by cyanogen bromide (Itakura et al. 1977; Goeddel et al. 1979; Nagai and Thøgersen 1984; Sung et al. 1986). The fusion protein is often stable compared to the foreign protein, and a protein can be made that does not contain an amino-terminal methionine.

Nagai and Thøgersen (1984) have developed a vector designed to produce hybrid proteins that can be cleaved with the blood coagulation factor X_a to liberate the foreign protein. Factor X_a cleaves specifically after the arginine of the tetrapeptide Ile-Glu-Gly-Arg, which connects the 31 amino-terminal residues of the bacteriophage λ *c*II protein to the protein of interest (Figure 17.10).

Other proteases have been similarly employed. For example, Jia et al. (1987) used λgt11 to make a β-galactosidase–carbohydrate-binding protein 35 (CBP35) fusion protein, which was then cleaved with V8 protease to release CBP35.

Proinsulin is readily degraded in *E. coli*. However, expression of a fusion protein composed of two molecules of proinsulin has been shown to stabilize the protein, perhaps because the protein product is larger than proinsulin. This fusion protein can then be cleaved by cyanogen bromide (Shen 1984).

In each case, it is important that the target amino acids for the protease or cyanogen bromide (which cleaves after methionine residues) does not occur at other locations within the protein of interest.

PRODUCTION OF HYBRID PROTEINS THAT CAN BE CLEAVED WITH FACTOR X_a

The expression plasmid of Nagai and Thøgersen is constructed as follows:

1. Using one of the three methods described on pages 17.18–17.24, prepare a cDNA fragment with a blunt 5′ terminus immediately preceding the first codon of the sequence to be expressed.

2. Clone the blunt-ended coding sequence into the *Stu*I site of M13mp11 FX. This *Stu*I site is located at the 3′ terminus of the sequence encoding the factor X_a recognition site.

3. Clone a *Bam*HI fragment encoding the factor X_a site and the foreign gene into the *Bam*HI site of pLcII, a plasmid encoding the bacteriophage λ p_L promoter that directs transcription of bacteriophage λ *c*II.

4. Transform *E. coli* strain MZ-1, which carries a defective bacteriophage λ *cIts*857 prophage, and plate on LB medium containing ampicillin (100 μg/ml). Incubate overnight at 30°C.

5. Test individual colonies for the presence of the desired insert by plasmid minipreps.

6. To obtain high levels of transcription, follow the procedure in step 5, page 17.11, for induction of the p_L promoter. Test correct clones for protein production (see pages 17.34–17.35).

7. Purify the fusion protein and cleave it with factor X_a as follows:

   a. Grow 500 ml of MZ-1 cells carrying the plasmid encoding the *c*II-Ile-Glu-Gly-Arg-foreign protein at 30°C in LB medium containing ampicillin (100 μg/ml) to an A_{600} of 0.7.

   b. Add 500 ml of LB medium preheated to 65°C, and continue incubation for 2 hours at 40°C.

   c. Collect the cells and purify the fusion protein (see Nagai and Thøgersen 1984). Store at −20°C until needed.

   d. Activate bovine blood coagulation factor X to factor X_a with Russell's viper venom (Fujikawa et al. 1972).

   e. Dissolve the purified fusion protein in 100 mM NaCl, 50 mM Tris · Cl (pH 8.0), 1 mM $CaCl_2$, and add factor X_a at a 1:100 (factor X_a: protein) molar ratio. Incubate for 2 hours at 25°C. The cleavage reaction should be tested by SDS-polyacrylamide gel electrophoresis. Depending on the use, the cleaved product should be either used directly or purified further.

Expression of Secreted Foreign Proteins

Secretion of the foreign protein is an alternative to intracellular expression. This is accomplished by fusing the coding sequence to DNA encoding a signal peptide that is cleaved by signal peptidase when the protein is secreted into the periplasm located between the inner and outer membranes of *E. coli*. The two major problems encountered in secretion of foreign proteins are that yields are often low and cleavage of the signal peptide may not occur or may occur at an inappropriate position. However, there are a number of advantages:

- Some proteins that are degraded by intracellular proteases are stable in the periplasm (Talmadge et al. 1980).
- Some proteins that are inactive when produced intracellularly are active when secreted; secretion may allow them to be properly folded (Gray et al. 1985).
- Proteins are produced that do not have an amino-terminal methionine, since cleavage occurs between the signal peptide and the coding sequence.

Several expression vectors have been designed for the secretion of foreign proteins (Talmadge et al. 1980; Oka et al. 1985; Takahara et al. 1985). It may be necessary to try more than one secretion system to find a vector and an *E. coli* strain that will allow secretion of a given foreign protein. Even then, processing may not occur in the correct position and protein levels may be low. One system that may allow secretion and processing takes advantage of the alkaline phosphatase promoter (*phoA*) and signal sequence. Upon phosphate starvation, this promoter and signal direct the synthesis of large amounts of alkaline phosphatase, which is secreted into the periplasm. This system has been used for secretion of human epidermal growth factor (hEGF) into the periplasm of *E. coli* (Oka et al. 1985). The secreted hEGF has the correct amino terminus.

Plasmid pTA1529 (Figure 17.11) is a convenient vector for *phoA*-mediated expression and secretion.

1. Using one of the three methods described on pages 17.18–17.24, prepare a DNA fragment encoding the foreign protein such that there is a blunt 5′ terminus immediately preceding the first codon of the sequence to be expressed.

2. Digest pTA1529 with *Hin*dIII and fill in recessed termini with the Klenow fragment of *E. coli* DNA polymerase I (see page 17.23, step 5a and b, and page 17.22, step 3g–h). Cleave the vector at an appropriate downstream site (*Eco*RI, *Sma*I, *Bam*HI).

3. Clone the fragment into the vector, transform *E. coli* strain YK537, and plate on LB medium containing ampicillin (100 μg/ml). Screen transformants with a junction oligonucleotide or by plasmid minipreps, and sequence the junction region of the potentially correct clones to verify (see Chapter 13).

4. To obtain high levels of transcription from the *phoA* promoter, grow an overnight culture of YK537 transformants in high-phosphate medium (640 μM KH_2PO_4). Dilute the culture 1:50 in low-phosphate medium (32 μM KH_2PO_4). Phosphate will gradually become limiting, and the *phoA* promoter will be induced. Test for protein production at various times (2, 4, 8, and 24 hours) after induction using one of the methods described on pages 17.34–17.35.

5. Determine the cellular location of the protein by cell fractionation.

   a. Place 1 ml each of the induced and uninduced culture in separate microfuge tubes. Centrifuge at 12,000*g* for 1 minute at 4°C in a microfuge to pellet the cells. Discard the supernatant.

   b. Suspend the cell pellets in 100 μl of a freshly prepared solution of lysozyme (1 mg/ml), 20% w/v sucrose, 30 mM Tris · Cl (pH 8.0), 1 mM EDTA (pH 8.0), and place on ice for 10 minutes.

      Lysozyme will not work efficiently if the pH of the solution is less than 8.0.

   c. Recover the cells by centrifugation as in step a. The supernatant is the periplasmic fraction, which should be stored at 4°C until step g.

   d. Resuspend the cell pellets in 0.1 M Tris · Cl (pH 8.0), and break open the cells by freezing and thawing (i.e., place the cells on dry ice, thaw at 37°C, and repeat).

   e. Centrifuge the suspension at 12,000*g* for 5 minutes at 4°C in a microfuge. The cytoplasmic proteins are mainly in the supernatant, which should be stored at 4°C until step g. The pellet contains the membrane fraction and insoluble inclusion bodies.

   f. Solubilize the membrane proteins by incubating the pellet in 1% Triton X-100 for 10 minutes at 4°C.

g. Analyze the periplasmic, cytoplasmic, and membrane proteins by SDS-polyacrylamide gel electrophoresis (see Chapter 18, pages 18.47–18.54), potentially coupled with western blotting (see Chapter 18), to determine how effectively the protein of interest is secreted into the periplasm.

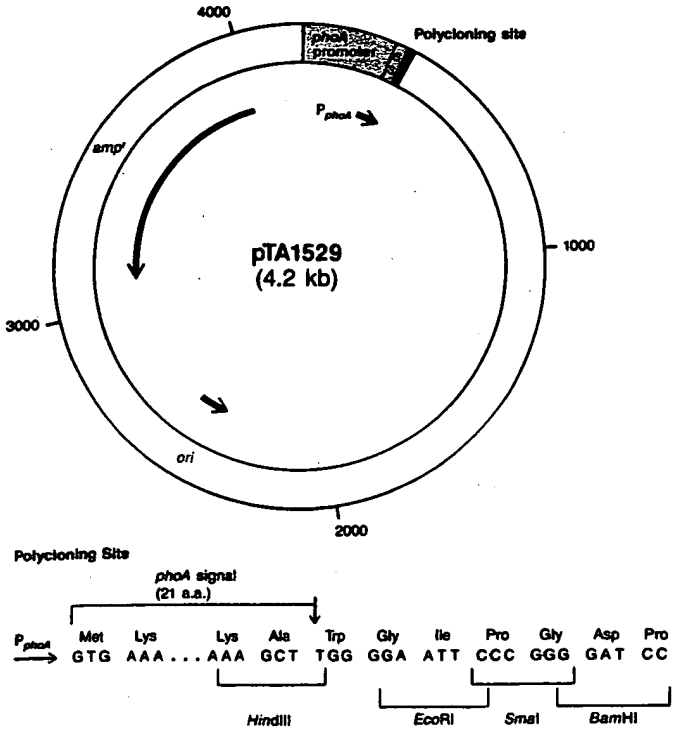


FIGURE 17.11
pTA1529 contains the promoter and signal sequence (SS) of the *E. coli phoA* gene. The first codon of the signal sequence is GTG (Kikuchi et al. 1981). A polycloning site is present near the signal sequence cleavage site (↓) (Oka et al. 1985).

QUANTITATING THE LEVELS OF EXPRESSION OF CLONED GENES

After plasmids containing the transcription and translation initiation signals that allow the protein of interest to be expressed in *E. coli* are constructed, the amount of protein made is measured. Generally, four approaches are used:

1. SDS-polyacrylamide gels are run to determine whether a protein of the appropriate size is made at increased levels only after induction in cells that contain the plasmid expressing the protein of interest (Laemmli 1970). Often the protein of interest can be visualized by staining the gel with Coomassie Brilliant Blue or by silver staining (Oakley et al. 1980) (see Chapter 18, pages 18.55–18.57). If no new protein band is seen using either of these stains, metabolic labeling with 100 μCi of [^{35}S]Met or [^{35}S]Cys per milliliter of culture for 5 minutes following induction (see Chapter 18, pages 18.27–18.28), followed by SDS-polyacrylamide gel electrophoresis and autoradiography, may allow detection of the protein of interest.

2. Western gel analysis (Towbin et al. 1979) (see Chapter 18, pages 18.60–18.75) using antibodies that bind specifically to the protein of interest is usually necessary to confirm the identity of the gene product detected by SDS-polyacrylamide gel electrophoresis. In addition, sometimes proteins that are not detectable in Laemmli gels are detected by western gels.

3. Activity assays are useful in evaluating the amount of active protein produced and, therefore, in evaluating any attempts to increase production. Sometimes activity does correlate with the level of protein produced. However, if the protein is made in an inactive form or is sequestered in inclusion bodies, then activity may not parallel the level of expression. In some cases, the goal is to make high levels of protein. However, if procedures are not established to convert inactive protein to an active form, then it may be prudent to optimize for activity.

4. If expression is low, it is often useful to place the *lacZ* gene downstream from the gene that is being expressed. This makes it possible to measure the amount of transcription that has proceeded through the gene. Thus, if transcription or translation is limiting expression, changes in the expression system (such as those described on page 17.36) can be monitored by changes in β-galactosidase activity (Miller 1972).

Monitoring Expression by β-Galactosidase Activity

1. Grow an overnight culture in A medium containing 0.4% glucose, 1 μg/ml vitamin B1, 1 mM $MgSO_4$, and the appropriate antibiotic. Dilute the culture 1:50 in the same medium and grow to an A_{600} of 0.4. Place the culture on ice.

   To prepare A medium, mix:

K_2HPO_4	10.5 g
KH_2PO_4	4.5 g
$(NH_4)_2SO_4$	1.0 g
Na citrate · $2H_2O$	0.5 g
H_2O to 1 liter	

2. Mix 0.1 ml of the cell culture, 0.9 ml of Z buffer, 2 drops of chloroform, and 1 drop of 0.1% SDS. Vortex for 10 seconds, and then equilibrate to 28°C.

Z buffer

0.06 M $Na_2HPO_4 \cdot 7H_2O$
0.04 M $NaH_2PO_4 \cdot H_2O$
0.01 M KCl
0.001 M $MgSO_4 \cdot 7H_2O$
0.05 M β-mercaptoethanol

The pH of the buffer should be 7.0. Do not autoclave.

3. Add 0.2 ml of *o*-nitrophenyl-β-D-galactopyranoside in A medium (4 mg/ml) to the lysed culture prepared in step 2 at 1-minute intervals. When a yellow color develops, stop the reaction by adding 0.5 ml of 1 M Na_2CO_3. Record the time at which the yellow color develops.

4. Determine the A_{600} of the cultures. Read the A_{550} and A_{420} of the reaction samples. Units of β-galactosidase activity =

$$1000 \times \frac{A_{420} - (1.75 \times A_{550})}{t \times 0.1 \times A_{600}}$$

where t = time in minutes. Improvements in transcription through the gene should be reflected by increases in the amount of β-galactosidase activity.

INCREASING EXPRESSION OF CLONED GENES

The level of expression of the cloned gene may be low due to RNA instability, premature termination, inefficient translation, or protein instability. Protein instability is distinguished from the other potential problems by pulse-chase experiments. A culture of cells carrying the expression plasmid is grown to mid-log phase, induced, and labeled for 10 minutes with [^{35}S]Met or [^{35}S]Cys (see Chapter 18, pages 18.27–18.28), followed by the addition of excess methionine or cysteine. Samples taken every 10 minutes from immediately before induction to 1 hour after induction are run on an SDS-polyacrylamide gel. The gel is dried and subjected to autoradiography to determine whether the protein is unstable. If the protein is unstable, then the use of protease-deficient strains (*lon*, *hfl*, or *htpR*) (Buell et al. 1985) or protease inhibitors during growth and harvesting may increase yields (1 mM phenylmethylsulfonyl fluoride [PMSF], 1–5 mM EDTA, 20 mM benzamidine). In many cases, problems of protein instability have been overcome by increasing the amount of foreign protein synthesized.

Several strategies have been successful in increasing the synthesis of foreign proteins. These include:

- *Use of site-directed mutagenesis to increase translation initiation.* The presence of secondary structure involving the ribosome-binding site can reduce the efficiency of translation initiation. Thus, removal of potential secondary structure involving the ribosome-binding site can increase translation initiation (Iserentant and Fiers 1980; Queen and Rosenberg 1981; Hall et al. 1982; Coleman et al. 1985). The SD sequence and the initiation codon are the major factors in ribosome binding. However, conserved base pairs in the region from −20 to +13 have been reported (Gold et al. 1981), and mutations in base pairs other than those in the SD sequence or ATG have been shown to alter translation initiation (Hui et al. 1984; Wood et al. 1984). Changing the DNA sequence near the SD and ATG sequences may increase expression.

- *Use of mutant strains defective in termination of transcription (*rho*) or altered in RNA metabolism (*pnp, rna*) to increase the amount of functional RNA.* Also try several different strains of *E. coli* because expression levels vary in different strains for reasons that are not clear. If the p_L promoter is used, then the strains must be lysogenized with bacteriophage λ *cIts*857. If the *tac* promoter is used, then F′ *lacI*^{q} must be transferred to the host or *lacI* must be cloned into the expression plasmid.

- *Placing sequences that may stabilize the mRNA or make it more translatable downstream from the inserted DNA (i.e., transcription terminators or RNAase III sites [Panayotatos and Truong 1981; Studier and Moffatt 1986; Rosenberg et al. 1987]).* Studies of sib mutants of bacteriophage λ demonstrate the effect of secondary structure on the stability of the mRNA in *E. coli.*

- *Altering the kinetics of synthesis by changing temperature, ribosome-binding site (Marquis et al. 1986), promoter strength, plasmid copy number, or amount of inducer.* The kinetics of synthesis of foreign proteins may change cell physiology, protein folding, or susceptibility to proteolysis.

PROTEIN PURIFICATION

Inclusion Bodies

A high level of expression of proteins in *E. coli* often results in cytoplasmic granules that can be seen with a phase-contrast microscope and that can be separated from crude cell lysates by centrifugation. Cells expressing high levels of foreign protein are concentrated by centrifugation and lysed by mechanical techniques, sonication, or lysozyme plus detergents. The inclusion bodies are pelleted by centrifugation and washed with Triton X-100 and EDTA (Marston et al. 1984) or urea (Schoner et al. 1985).

To obtain soluble, active protein, the washed inclusion bodies must be solubilized and then refolded. Each protein may require a different procedure, which must be determined empirically. Various conditions (e.g., guanidine HCl [5–8 M], urea [6–8 M], SDS, alkaline pH, or acetonitrile/propanol) may be used to solubilize the inclusion bodies. A number of strategies to identify effective solubilization reagents have been described by Marston (1987). The procedure given below has been used to solubilize prorennin inclusion bodies and is adapted from Marston et al. (1984).

After successful solubilization, various refolding methods involving dilution or dialysis may be tried. The yield of active protein or of protein with the same disulfide bonds as the original protein depends on the concentration, purity, and size of the polypeptide; the pH and ionic strength of the solvent; and the rate of refolding. Other factors include the number of disulfide bonds and the nature of the protein itself. A number of successful refolding protocols are detailed in Marston (1987).

Takahara, M., D.W. Hibler, P.J. Barr, J.A. Gerlt, and M. Inouye. 1985. The *ompA* signal peptide directed secretion of staphylococcal nuclease A by *Escherichia coli*. *J. Biol. Chem.* **260:** 2670.

Talmadge, K., J. Kaufman, and W. Gilbert. 1980. Bacteria mature preproinsulin to proinsulin. *Proc. Natl. Acad. Sci.* **77:** 3988.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc. Natl. Acad. Sci.* **76:** 4350.

Weinstock, G.M., C. ap Rhys, M.L. Berman, B. Hampar, D. Jackson, T.J. Silhavy, J. Weisemann, and M. Zweig. 1983. Open reading frame expression vectors: A general method for antigen production in *Escherichia coli* using protein fusions to β-galactosidase. *Proc. Natl. Acad. Sci.* **80:** 4432.

Wood, C.R., M.A. B ss, T.P. Patel, and J.S. Emtage. 1984. The influence of messenger RNA secondary structure on expression of an immunoglobulin heavy chain in *Escherichia coli*. *Nucleic Acids Res.* **12:** 3937.